Space non-invariant point-spread function and its estimation in fluorescence microscopy

Praveen Pankajakshan, Laure Blanc-Féraud, Zvi Kam, Josiane Zerubia

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Space non-invariant point-spread function and its estimation in fluorescence microscopy.

Praveen PANKAJAKSHAN, Laure BLANC-FÉRAUD, Zvi KAM and Josiane ZERUBIA

N° 7157

December 2009

Thème COG

ISSN 0249-6399 ISRN INRIA/RR--7157--FR+ENG

Space non-invariant point-spread function and its estimation in fluorescence microscopy.

Praveen PANKAJAKSHAN*, Laure BLANC-FÉRAUD*, Zvi KAM† and Josiane ZERUBIA*

Thème COG — Systèmes cognitifs
Projet ARIANA

Rapport de recherche n° 7157 — December 2009 — 56 pages

Abstract: In this research report, we recall briefly how the diffraction-limited nature of an optical microscope's objective, and the intrinsic noise can affect the observed images' resolution. A blind deconvolution algorithm can restore the lost frequencies beyond the diffraction limit. However, under other imaging conditions, the approximation of aberration-free imaging, is not applicable, and the phase aberrations of the emerging wavefront from a specimen immersion medium cannot be ignored any more. We show that an object's location and its original intensity distribution can be recovered by retrieving the refracted wavefront's phase from the observed intensity images. We demonstrate this by retrieving the point-spread function from an imaged microsphere. The noise and the influence of the microsphere size can be mitigated and sometimes completely removed from the observed images by using a maximum a posteriori estimate. However, due to the incoherent nature of the acquisition system, phase retrieval from the observed intensities will be possible only if the phase is constrained. We have used geometrical optics to model the phase of the refracted wavefront, and tested the algorithm on some simulated images.

Key-words: fluorescence microscopy, point-spread function, blind deconvolution, spherical aberration, maximum a posteriori, maximum likelihood, expectation maximization

* ARIANA Project-team, INRIA/I3S/UNS, 2004 Route des lucioles-BP93, 06902 Sophia-Antipolis Cedex, France
† Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel 76100.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

La fonction de flou sans invariance spatiale et son estimation pour la microscopie à fluorescence.

Résumé : Dans la première partie de ce rapport de recherche, nous rappelons brièvement comment la nature limitée de diffraction de l'objectif d'un microscope optique, et le bruit intrinsèque peuvent affecter la résolution d'une image observée. Un algorithme de déconvolution aveugle a été proposé en vue de restaurer les fréquences manquants au delà de la limite de diffraction. Cependant, sous d'autres conditions, l'approximation du systéme imageur l'imagerie sans aberration n'est plus valide et donc les aberrations de la phase du front d'onde émergeant d'un médium ne sont plus ignorées. Dans la deuxième partie de ce rapport de recherche, nous montrons que la distribution d'intensité originelle et la localisation d'un objet peuvent être retrouvées uniquement en obtenant de la phase du front d'onde réfracté, à partir d'images d'intensité observées. Nous démontrons cela par obtention de la fonction de flou a partir d'une microsphère imagée. Le bruit et l'influence de la taille de la microsphère peuvent être diminués et parfois complètement supprimes des images observées en utilisant un estimateur maximum a posteriori. Néanmoins, a cause de l'incohérence du système d'acquisition, une récupération de phase a partir d'intensités observées n'est possible que si la restauration de la phase est contrainte. Nous avons utilisé l'optique géométrique pour modéliser la phase du front d'onde réfracté, et nous avons teste l'algorithme sur des image simulées.

Mots-clés : microscopie à fluorescence, fonction de flou, déconvolution aveugle, aberration sphérique, maximum a posteriori, maximum de vraisemblance, espérance-maximisation.

Acknowlegement

This research was funded by the P2R Franco-Israeli Collaborative Research Program[1]. We would like to thank INRIA for supporting the PhD of the first author through a CORDI fellowship. The authors gratefully acknowledge Dr. Bo Zhang (Philips Medical research), Prof. Olivier Haeberlé and Prof. Alain Dieterlen (Université de Haute-Alsace, Mulhouse), and Dr. Jean-Christophe Olivo-Marin (Pasteur Institute, Paris) for several interesting discussions. Additionally, our sincere gratitude goes to Dr. Gilbert Engler (INRA Sophia-Antipolis) for the images presented in Figs. 20 and 21 of Appendix D.

[1]See http://www-sop.inria.fr/ariana/Projets/P2R/index.html for more information

Contents

List of Figures

Abbreviations

3D	Three dimension
2D	Two dimension
AFP	Actual focal position
AU	Airy units
CCD	Charge-coupled device
CG	Conjugate gradient
CLSM	Confocal laser scanning microscope
EM	Expectation maximization
FWHM	Full-width at half maximum
GFP	Green fluorescent protein
iff	if and only if
LSI	Linear space invariant
LSM	Laser scanning microscope
LSNI	Linear space noninvariant
MAP	Maximum a posteriori
ML	Maximum likelihood
MLE	Maximum likelihood estimate
MSE	Mean squared error
NA	Numerical aperture
NFP	Nominal focal position
OTF	Optical transfer function
pdf	Probability density function
PMT	Photomultiplier tube
PSF	Point-spread function
SNR	Signal to noise ratio
SA	Spherical aberration
WFM	Widefield microscope
w.r.t	with respect to

Notations

$(\cdot * \cdot)$	Linear space invariant convolution
$(\cdot)^*$	Complex conjugation operation
$\widehat{(\cdot)}$	Estimate
α	Semi-aperture angle of the objective lens
$A(\cdot)$	Apodization function
b	Background signal
$\mathcal{B}$	Fourier-Bessel transform or Hankel transform of zero-order
D	Diameter of the CLSM pinhole
$\delta(\cdot)$	Dirac-delta function
Δ_{xy}	Radial sampling size
Δ_z	Axial sampling size
$\mathbf{E}(\cdot,\cdot)$	Time-varying electric field
ϵ	Algorithm convergence factor
$\epsilon_{(\cdot)}$	Permitivity of a medium
$\mathcal{F}(\cdot)$	Fourier transform
$\mathcal{F}^{-1}(\cdot)$	Inverse Fourier transform
γ	Reciprocal of the photon conversion factor
h	Point-spread function (PSF) of an imaging system
$\mathbf{H}(\cdot,\cdot)$	Time-varying magnetic field
$i(\cdot)$	Observed image
j	Imaginary unit of a complex number
J_0	Bessel function of the first kind of order zero
$\mathbf{k}$	Vector of coordinates in the frequency space
λ	Average photon flux
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
$\mathcal{J}(\cdot)$	Energy function
(μ, σ^2)	Mean and variance of a normal distribution
$\mu_{(\cdot)}$	Permeability of a medium
n	Index of iteration for the estimation algorithm $\in \mathbb{N}^+$
$n_{(med)}$	Refractive index of a medium
$\mathbf{n}$	Outward normal to a surface
$\nabla(\cdot)$	Gradient of a vector field
$\nabla^2(\cdot)$	Laplacian of a scalar field
$\mathcal{N}_g(\cdot)$	Additive Gaussian noise
$\mathcal{N}_p(\cdot)$	Poisson distribution
(N_x, N_y)	Number of pixels in the radial plane
(N_z)	Number of axial slices or planes
$o(\cdot)$	Specimen or object imaged

Continued on Next Page...

$\boldsymbol{\omega}_{(\cdot)}$	Parameter vector to be estimated
Ω_s	Discrete spatial domain
Ω_f	Discrete frequency domain
$p(\cdot)$	Probability density function
$Pr(\cdot\|\cdot)$	Conditional probability
P	Pupil function
(ρ, ϕ, z)	Cylindrical coordinates
(ρ, ϕ, θ)	Spherical coordinates
Σ	Diffracting aperture
Θ	Parameter space ($\in \mathbb{R}$)
φ	Optical phase difference
φ_d	Defocus phase term
φ_a	Aberration phase term
$\mathbf{x}$	Vector of coordinates in object/image space (2D or 3D) ($\in \mathbb{R}^2$ or $\in \mathbb{R}^3$)
z	Axial distance in the image space ($\in \mathbb{R}$)

1 Background

Fluorescent microscopes [1, 2] use a highly focused laser spot to scan biological specimens in three dimension (3D) and to obtain optical image sections of the volume of interest. The specimen is treated with a dye such as the green fluorescent protein (GFP) and it fluoresce on excitation by an incident laser beam. By changing the objective to focus at different depths inside the specimen, and by collecting the emitted fluorescence at each plane, one can visualize the cells, tissues and embryos in 3D. In Fig. 1, we see a simple schematic of a confocal laser scanning microscope (CLSM) where the emission field energy is collected by placing a photomultiplier tube (PMT) at the position of the emission beam focus. The difference between the classical fluorescent microscopes such as widefield microscope (WFM) is that in the CLSM a pinhole is added before the detection stage. This pinhole restricts the total amount of light collected to the plane that is in focus (as shown by solid line in the schematic; dotted line represents the out of focus planes). The major advantage of

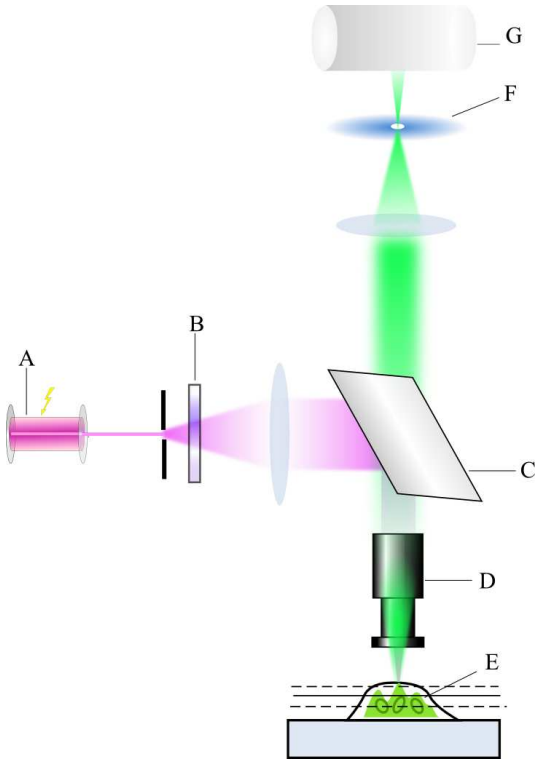


Figure 1: Schematic of a CLSM. A) Laser, B) excitation filter, C) dichromatic mirror, D) objective lens, E) in-focus plane of the specimen, F) pinhole aperture, G) photomultiplier tube (PMT) (© Ariana-INRIA/I3S).

using CLSM for imaging biological specimens is that it images efficiently those parts of the specimen that lie in the focal region of the excitation light. The light from the layers outside this region is greatly attenuated and does not contribute significantly to the final output. However, even under ideal conditions, the resolution of the observed object is affected by the diffraction-limited nature of the optical system. This is because when light from a point source passes through a small circular aperture, it does not produce a bright dot as an image, but rather a diffused circular disc known as Airy disc surrounded by much fainter concentric circular rings (see Fig. 2). This example of diffraction is of great importance because many

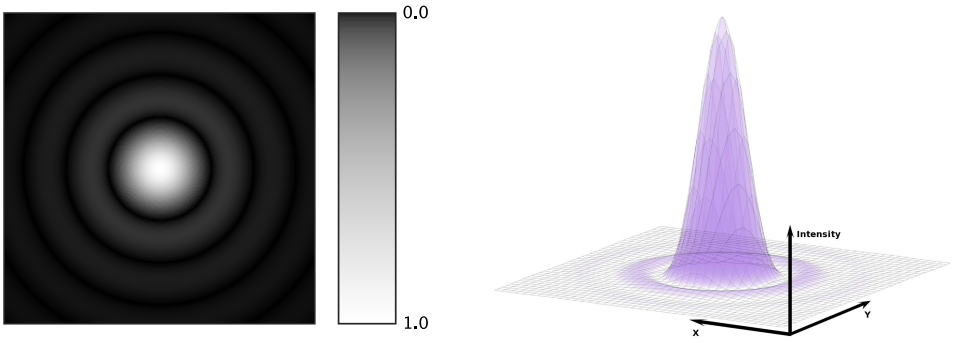


Figure 2: An illustration of the diffraction pattern for a circular aperture (© Ariana-INRIA/I3S).

optical instruments (including the human eye) have circular apertures. If this smearing of the image of the point source is larger than that produced by the aberrations of the system, the imaging process is said to be diffraction-limited, and that is the best resolution which can be physically obtained from that size of aperture.

In addition to blurring, the image measurement is corrupted by both intrinsic and extrinsic noise sources. In digital microscopy, the source of noise is either the signal itself (so-called photon shot noise), or the digital imaging system. As imaging is made possible by the conversion of light energy to photo-electrons at the detector element of the microscope, by tracking the accumulated photon to electron conversions (photo-electrons) alone over time, we can observe that it reveals an underlying Poisson distribution of events [3]. Thus, photon emission has the fundamental property of being stochastic with respect to time. On the other hand, if an image is taken with no light provided to the detector, the signal-independent electrons collect and form a distribution characterized by a mean value and a standard deviation. It is for this reason that the signal dependent noise is characterized by a Poisson distribution while imaging system noise usually follows a Gaussian distribution.

1.1 Imaging Model

Let $\mathcal{O}(\Omega_s) = \{o = (o_{xyz}) : \Omega_s \subset \mathbb{N}^3 \mapsto \mathbb{R}\}$ denote all possible observable objects on the discrete spatial domain $\Omega_s = \{(x, y, z) : 0 \leq x \leq N_x - 1, 0 \leq y \leq N_y - 1, 0 \leq z \leq N_z - 1\}$ and $h : \Omega_s \mapsto \mathbb{R}$ the microscope point spread function (PSF). If $\{i(\mathbf{x}) : \mathbf{x} \in \Omega_s\}$ (assumed to be bounded and positive) denote the observed volume, then the observation can be expressed as

$$i(\mathbf{x}) = \mathcal{N}_p([h * o](\mathbf{x}) + b(\mathbf{x})) + \mathcal{N}_g(\mathbf{x}),\ \mathbf{x} \in \Omega_s, \tag{1}$$

where, $\mathcal{N}_p(\cdot)$ denotes voxel-wise noise function modeled as a Poissonian process, $\mathcal{N}_g$ is the detector noise approximated by additive zero-mean Gaussian distribution $(0, \sigma_g^2)$, and $b : \Omega_s \mapsto \mathbb{R}$ denotes the low-frequency background signal caused by scattered photons and auto-fluorescence from the sample. Since the PMT in a confocal microscope (like Zeiss LSM 510 used in our experiments) usually has either a 8-bit or 12-bit sampling, the intensity gray levels are assumed to lie either between [0, 255] or between [0, 4095].

1.1.1 Background Fluorescence

In Eq. (1), we have assumed that the imaging system has been a priori calibrated so that there is negligible offset in the detector and that the illumination is uniform; that is no misalignment in the illumination lamp. This assumption is justified in our case as elucidated by the following example.

Fig. 3(a) and (b) show the first and the last slice of a real observed volume for a microsphere immersed in water. Fig. 3(c) and Fig. 3(d) show the estimated background $\hat{b}(\mathbf{x})$ obtained by morphologically opening the first (Fig. 3(a)) and the last (Fig. 3(b)) slices using a circular structural element [4]. The first and the last slices were chosen because they were found to be free of the object fluorescence. We can see from the histograms that there is a dominance of a single intensity value. Hence, the background is almost uniform and for our model we can assume it to be a constant. What we can also infer from the figures is that there is a uniformity in the illumination and no alignment problems. The mean value of the background signal from this estimation procedure was found to be about 3.13 intensity level units (IU). Fig. 4(a) shows the histogram calculated for the slice in Fig. 3(b). In the histogram shown in Fig. 4(b) for the background subtracted slice in Fig. 3(d), we notice now that the dominant intensities are essentially zero due to the subtraction. This shows that this kind of procedure works well in estimating the background. Next, the overall histogram of the image volume was smoothed and a Poisson distribution was fit to the data. The parameters of the distribution were estimated by using a maximum likelihood (ML) algorithm. The empirical mean of the distribution was estimated to be between 3.9683 and 3.9702 IU with a 95% confidence. Since the object fluorescence was sparsely populated, we find that there is not much difference between the overall distribution mean estimation and the individual background estimated. This is valid in most of the images taken using a CLSM where the object fluorescence is sparsely distributed throughout the volume. For more details on background estimation in fluorescence microscopy, the interested reader might refer to the following paper [5] by Van Kempen.

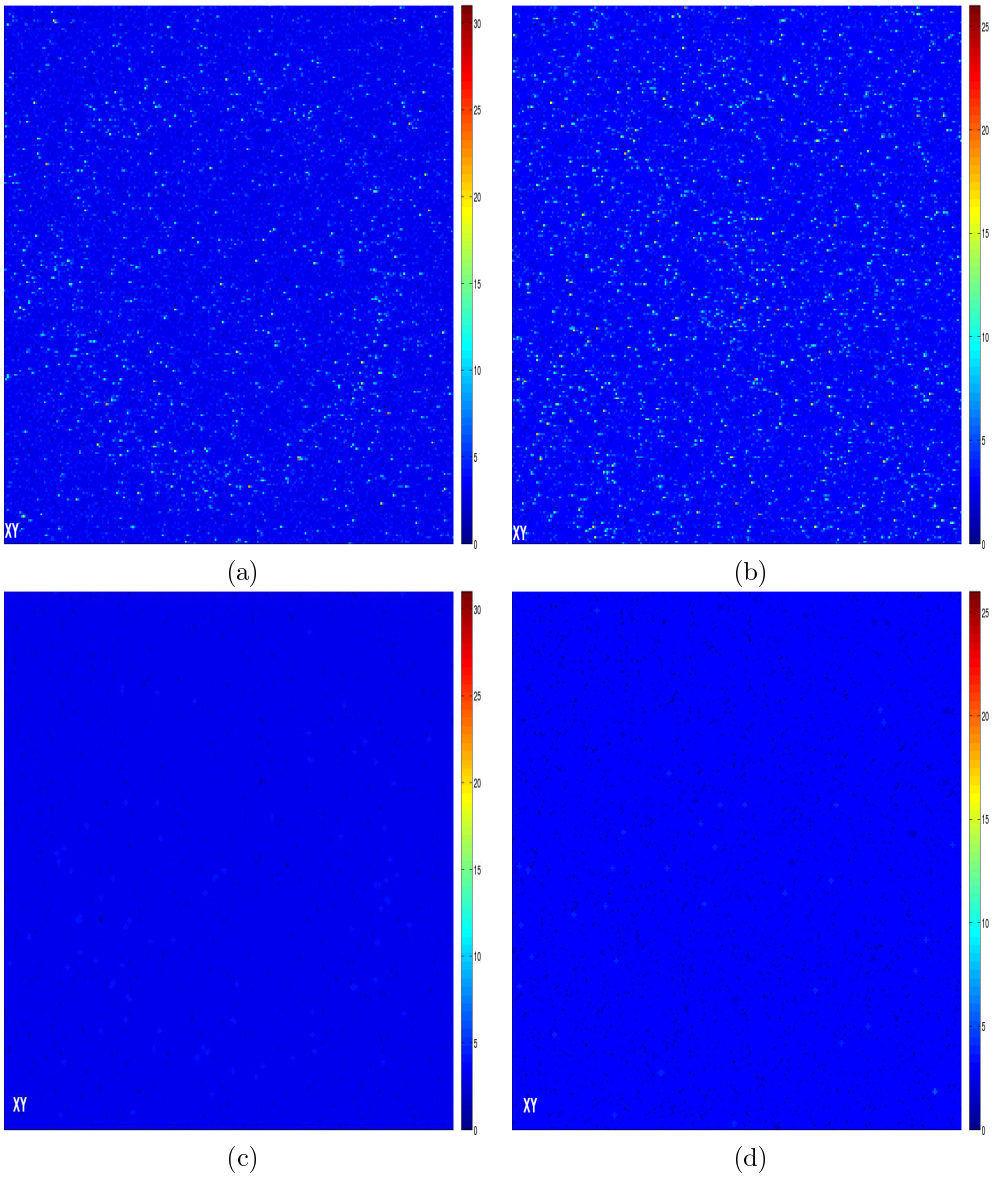
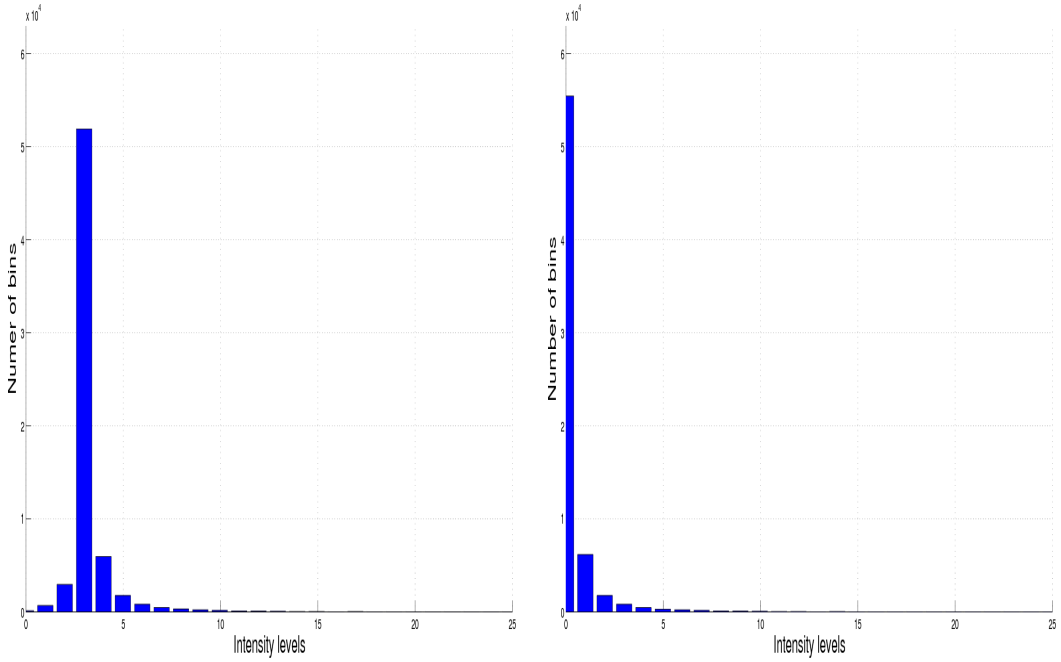


Figure 3: (a) The first and (b) the last slices of an observed microsphere image volume having sparse object fluorescence and the corresponding estimated background fluorescence $b(\mathbf{x})$ for (c) first and (d) last slices (© Ariana-INRIA/I3S).



(a) (b)

Figure 4: Histogram of (a) a slice in the original volume, and (b) the background subtracted slice (© Ariana-INRIA/I3S).

For the simulations used in the rest of this research report, a constant background fluorescence of intensity 10 was added to the blurred object and this is illustrated in Fig. 5(b). We can note the shift in axial position between the original object in Fig. 5(a) and its blurred version in Fig. 5(b). This axial shift will be further discussed in Subsection 2.2.1. It is important to notice that the radial center for the object and the observation remain unchanged.

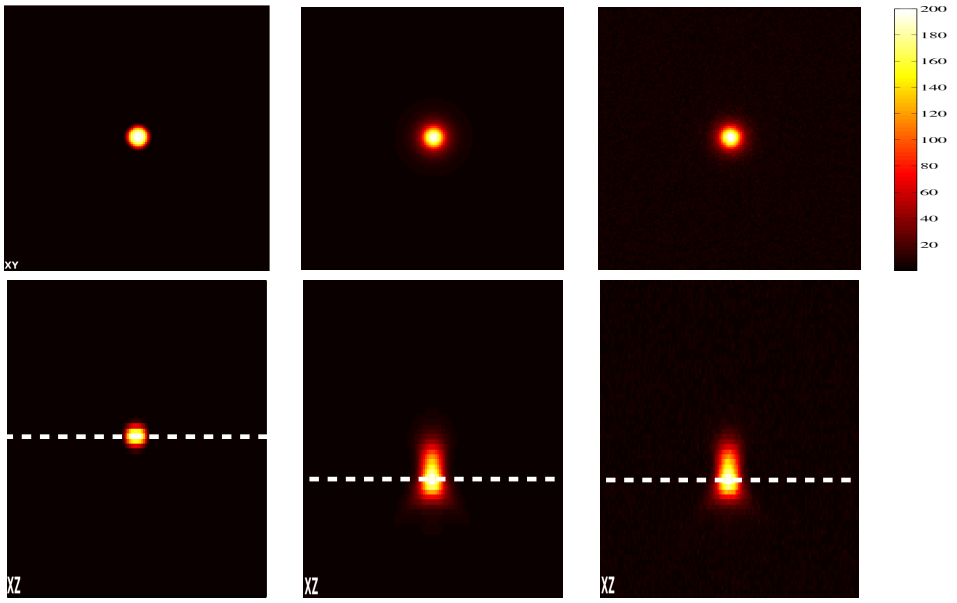
1.1.2 Noise

For a PMT operating in the photon-counting mode, we can assume that there is no *readout* or *quantum* noise, and the total photons detected due to object fluorescence and various other noise sources are Poisson distributed. Thus, the simulated data set i can be obtained as a Poisson random variable with mean $([h * o](\mathbf{x}) + b(\mathbf{x}))$ and Eq. (1) can be rewritten in terms of the observed photon counts as

$$\gamma i(\mathbf{x}) = \gamma(\mathcal{N}_p([h * o](\mathbf{x}) + b(\mathbf{x}))), \mathbf{x} \in \Omega_s, \tag{2}$$

where $\gamma \in \mathbb{R}^+$ is the reciprocal of the photon conversion factor, and $\gamma i(\mathbf{x})$ is the observed photon number at the detector. For fluorescence microscopy, γ is proportional to several physical parameters, such as the integration time and the quantum efficiency of the detector. It is important to note that the linear shift invariant (LSI) approximation for the convolution operation in the model of Eq. (2) is not valid when handling space varying PSFs. In this case, we will have to provide a new expression that is based on a new linear depth varying (LDV) observation model.

Although the Poisson statistics provide a realistic noise model, in many applications additive Gaussian noise model is assumed to simplify the numerical computation. However, it is important to note that under low signal to noise ratio (SNR), the additive Gaussian noise model provides a poorer description of the fluorescence microscopy imaging than the Poisson model. The high SNR case can be addressed by employing the central limit theorem (CLT) for large numbers of measurement data, where the additive Gaussian noise model performs satisfactorily. In Fig. 5(c), the microsphere of Fig. 5(a) is simulated as observed by a Zeiss LSM 510 microscope with a 40X oil immersion objective lens of numerical aperture (NA) 1.3.



(a) (b) (c)

Figure 5: Simulation of a (a) microsphere object, (b) the blurred observation showing the background fluorescence, (c) the blurred observation with the background fluorescence and Poisson noise with $\gamma = 100$ (© Ariana-INRIA/I3S).

1.1.3 Simulating Band-Limited Object

Since we frequently use simulated data set for testing or comparing the performance of our algorithm, it is necessary to numerically approximate a band-limited object that could be used to represent a real object that we are interested in imaging. The geometry of the object we wish to scan here, namely the microbead, is a sphere. Such spheres serve well as a good model for fluorescence-coated beads.

These spheres can be generated from their analytical expressions in the frequency domain as [6]

$$O(\mathbf{k}) = \frac{\sin(2\pi R\sqrt{k_x^2+k_y^2+k_z^2}) - (2\pi R\sqrt{k_x^2+k_y^2+k_z^2})\cos(2\pi R\sqrt{k_x^2+k_y^2+k_z^2})}{\pi^2(\sqrt{k_x^2+k_y^2+k_z^2})^3} \tag{3}$$

where R is the radius of the desired sphere, and the sampling in the frequency domain is carried out so that $\Delta_{k_x} = 1/N_x$, $\Delta_{k_y} = 1/N_y$, and $\Delta_{k_z} = \Delta_{xy}/(\Delta_z N_z)$. Δ_{xy} and Δ_z are respectively the radial and the axial sampling in the image space. To avoid Gibbs phenomenon, the numerical approximation of $O(k_x, k_y, k_z)$ was multiplied by a ℓ_∞ normalized 3D Gaussian function with variances $(N_x, N_y, N_z)/(2\pi)$. Fig. 5(a) shows the simulated sphere with manufacture specified radius of 250nm. The object intensity is assumed to lie between [0, 200].

The simulation is performed as if the microsphere images were acquire for a Zeiss LSM 510 microscope equipped with an Argon ion scanning laser of wavelength 488nm. The emission Band Pass (BP) filters are BP $505-530$ for a Green Fluorescence Protein (GFP) type staining.

1.2 Problem Statement

The fluorescence distribution in an object can be restored by deconvolution with the knowledge of the PSF of the imaging system. It has been widely accepted in the literature that the deconvolution performance of the Gaussian noise model is inferior to the physically correct Poisson noise model. We thus justify the use of the model in Eq. (2). We observed that theoretically calculated PSFs often lack the experimental or microscope specific signatures, while the empirically obtained images of microspheres are either over sized or (and) too noisy. Fig. 6 shows the axial intensity profiles along a radial plane for a numerically calculated PSF, a simulated object blurred with the PSF and a simulated observation. As we deal with live cells, often only a single observation of the volume is available for restoration. This is because, the very plant and animal cells that we wish to observe could be irreversibly damaged when exposed to high energy radiation for a long time. Photobleaching of the dye is another reason for avoiding long exposures. The non-repeatability of the whole process makes it difficult to restore the lost frequencies beyond the diffraction limits without any constraints on the object or the PSF. In the first phase of our work [7, 8, 9, 10], we proposed a blind deconvolution algorithm that estimated the diffraction-limited PSF (under aberration free conditions) and alternatively iterated to restore the true object fluorescence. We

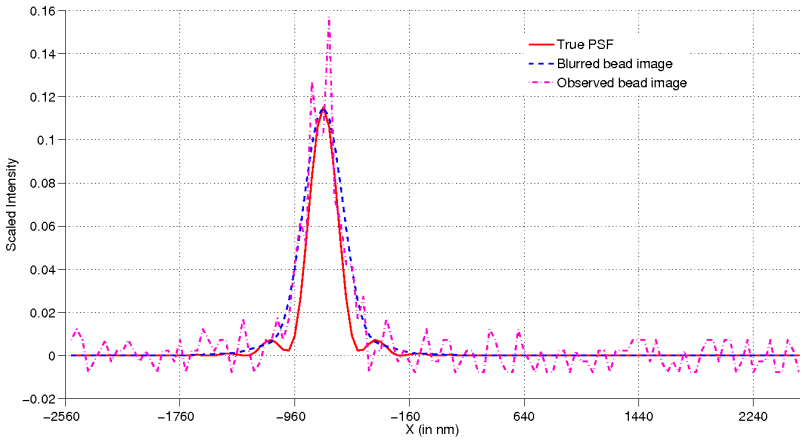


Figure 6: Axial intensity profiles of a calculated PSF, a blurred microsphere and the observed microsphere. The intensities are scaled and the peaks matched for visual comparison (© Ariana-INRIA/I3S).

minimized an energy function alternatively, first with respect to the intensity of an imaged object and then with respect to the PSF of the microscope. At each iteration the estimated object intensity was regularized using a total variational potential function, and the PSF was projected onto a parametric space. We projected the estimated PSF on to a 3D Gaussian model [11], and restored the lost frequencies beyond the diffraction limits by regularizing the object using a total variational (TV) term.

In most cases the degradation is often influenced only by the diffraction effect. However, if imaging into deeper sections of the specimen, spherical aberrations (SA) cannot be ignored. This is because, the refractive index mismatch between the specimen and the immersion medium of the objective lens becomes significant [12]. An additional difference in the path is introduced in the emerging wavefront of the light due to this difference in the index. The aim of this report is to derive a methodology for retrieving the true PSF of the imaging system from the observed images of microspheres in the presence of aberrations. Henceforth, we shall address the images of microspheres as experimental PSF. It is the usual procedure to use this experimental PSF for deconvolving a separate set of specimen observation data. In [13], it was noticed that an iterative nonlinear deconvolution algorithm when applied is very sensitive to the amount of randomness in the experimental PSF. Moreover an estimation procedure carried out on the observational data showed that the recovered objects are significantly improved by using a smoothed experimental PSF rather than using the experimental PSF directly. Similarly, for a deterministic algorithm tested on simulated objects, the root mean squared error (RMSE) is significantly smaller for the smoothed experimental PSF than using the non-smoothed one. An additional point that has not been addressed before, which we would like to point out, is that the microspheres in such a case

are too large to represent an ideal point source. In contrast, if the size of the beads are too small, then the observation is plagued with noise problems as the bead is weakly fluorescent. We summarize the motivation for the research in this report as follows

- experimentally obtained PSFs by imaging fluorescent beads are low in contrast and highly noisy,
- they are large in size compared to the actual PSF,
- the PSFs so constructed cannot be reused if the experimental conditions are varied, and
- knowledge of the PSF of an imaging system helps in recovering the original intensity distribution of an imaged object by deconvolution.

1.3 Organization of this Research Report

This research report is organized as follows. In Section 2, we retrace the scalar diffraction theory for finally introducing the Stokseth's [14] approximation for the diffraction-limited analytical PSF expression. This model is extended in Subsection 2.2 to also include aberrations. In Section 3, we enlist the existing literature on the subject of blind deconvolution and we see why these methods cannot be applied for solving the above problem. Drawing inspiration from our earlier work on the subject of blind deconvolution, we propose in Section 4 an approach of estimating the actual PSF from the observation data with Bayesian framework. The difference here is that we are more interested in estimating PSFs that vary with the acquisition conditions; but with some knowledge of the object. To test our hypothesis, in Section 5, we have shown some experiments performed on simulated data. Finally we discuss the possibility of extending this work to retrieve the PSF directly from observed specimen data and finally we conclude in Section 6.

2 Analytical Point-Spread Function Model

In this section we review some of the basics of the diffraction theory that will help us in eventually deriving an analytical expression for the PSF. Although electromagnetic fields are vectorial in nature, we have considered only the scalar properties of light here. Our reasons for ignoring the vectorial nature of light will be made clear in the discussion at the end of this section.

2.1 Review of the Scalar Diffraction Theory

2.1.1 Maxwell and Helmholtz's Equations

In the absence of free charge, Maxwell's equations are [15]

$$\nabla \times \mathbf{E} = -\mu_m \frac{\partial}{\partial t}\mathbf{H}, \tag{4}$$

$$\nabla \times \mathbf{H} = \epsilon_m \frac{\partial}{\partial t}\mathbf{E}, \tag{5}$$

$$\nabla \cdot \epsilon_m \mathbf{E} = 0, \tag{6}$$

$$\nabla \cdot \mu_m \mathbf{H} = 0, \tag{7}$$

where $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ are the orthogonally time-varying electric and magnetic fields respectively (see Fig. 7). $\nabla \times \mathbf{E}$ gives the curl($\mathbf{E}$) and $\nabla \cdot \epsilon_m \mathbf{E}$ gives the div($\epsilon_m \mathbf{E}$). μ_m is the permeability and ϵ_m is the permittivity of the medium. If the medium is homogeneous, ϵ is constant throughout the region of propagation. The medium is said to be non-dispersive to light if ϵ is independent of the wavelength λ of the light used. All media are non-magnetic hence the permeability of the medium is the same as that in vacuum. The vector wave

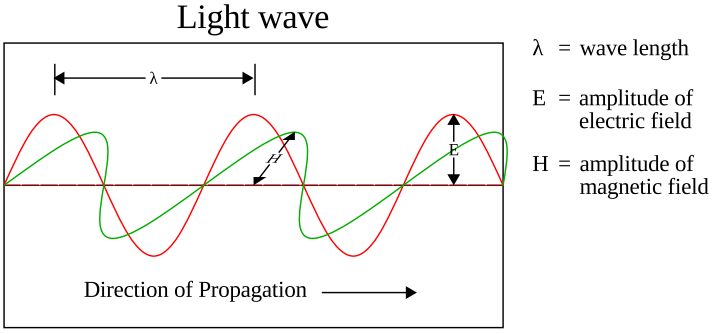


Figure 7: Light as an electromagnetic wave (© Ariana-INRIA/I3S).

equation for any time-varying field $\mathbf{V}(t)$ can thus be written as

$$\nabla^2\mathbf{V}(t) - \frac{n^2}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{V}(t) = 0, \tag{8}$$

where, ∇^2 is the Laplacian operator, $n_m = (\epsilon_m/\epsilon_0)^{1/2}$ is the refractive index of the medium, and $c = 1/(\mu_0\epsilon_0)^{1/2}$ the speed of light in vacuum. However, for non homogeneous medium, the coupling between the electric and magnetic fields cannot be rejected and hence the above equation has to be modified to include also the variation in the refractive index. At the boundaries, coupling is introduced between the electric and magnetic fields and in their individual components as well. The error is small provided that the boundary conditions have effect over an area that is a small part of the area through which a wave may be passing.

We define a strictly monochromatic time-harmonic scalar field by

$$U(\mathbf{x}, t) = U_{\mathbf{x}}(\mathbf{x})\exp(-j\omega t). \tag{9}$$

The above scalar field also satisfies Eq. (8), and hence the Helmholtz equation

$$(\nabla^2 + k^2)U = 0, \tag{10}$$

where $k = (2\pi n\mu)/c = 2\pi n/\lambda$ is the wave number, and λ is the wavelength of the light in free space ($\lambda = c/\mu$). If this field has no evanescent component, it can be written as a weighted sum of plane waves of the form

$$U(\mathbf{x}) = \int_{\mathbf{k}} A(\mathbf{k})\exp(j\mathbf{k}\cdot\mathbf{x})d\Omega, \tag{11}$$

where $\mathbf{k}$ is a unit vector that describes the direction of propagation of the plane waves.

2.1.2 Fundamentals of Vector Calculus

Divergence Theorem The Divergence theorem or the Gauss' theorem is the higher dimensional form of the fundamental theorem of calculus.

Theorem 2.1. *Let V be a simple solid region and S the boundary surface of V, given with the positive outward orientation* $\mathbf{n}$. *Let* $\mathbf{V}$ *be a vector field whose component functions have continuous partial derivatives on an open region that contains V. Then*

$$\int_V (\nabla\cdot\mathbf{V})dV = \oint_S \mathbf{V}\cdot\mathbf{n}dS. \tag{12}$$

It relates the flux of a vector field through a surface S to its behavior inside the surface.

Green's second identity Green's identities are a set of vector derivative/integral identities that can be useful in deriving the Fresnel-Kirchoff diffraction equations and the Debye integral approximation. Since we are only interested in the second identity, it is stated as follows.
If G and U are continuously differentiable scalar fields on V in $\mathbb{R}^3$, then

$$\int_V (G\nabla^2 U - U\nabla^2 G)dV = \oint_S (G\frac{\partial}{\partial \mathbf{n}}U - U\frac{\partial}{\partial \mathbf{n}}G) \cdot dS. \tag{13}$$

Here, $G(P)$ is a scalar field as a function of position P, and $\partial/\partial\mathbf{n}$ is the partial derivative along the outward normal direction (see Fig. 8) in the surface element dS. It is straightforward to show that for the scalar fields G and U satisfying the Helmholtz equation Eq. (10), the left hand side of Green's second identity Eq. (13) is zero.

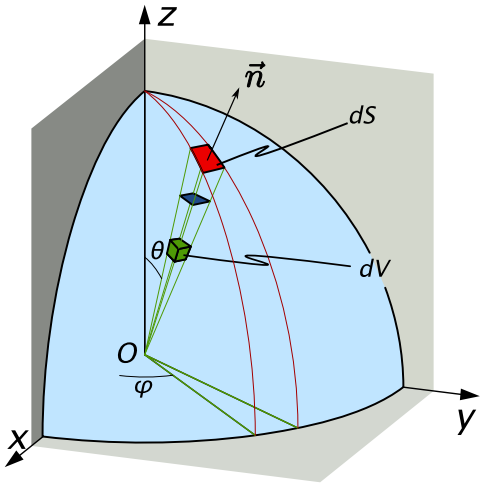


Figure 8: Solid region depiction for Green's identity and the divergence theorem (© Ariana-INRIA/I3S).

2.1.3 Scalar Diffraction Theory

We introduce below the diffraction theory for light propagation in a homogeneous medium. The most important approximation here is the treatment of light as a scalar phenomenon, neglecting the fundamentally vectorial nature of the electromagnetic fields. The scalar theory yields accurate results if two conditions are met:

1. the diffraction aperture Σ is very large in comparison to the wavelength of light λ, and
2. the diffracting fields must not be observed too close to the aperture $r_{01} \gg \lambda$.

For all microscopes operating in the far-field region, the above approximations are justified. In Fig. 9, if $\mathcal{P}_0$ is the point of observation, for diffraction by an aperture Σ, the Kirchoff G

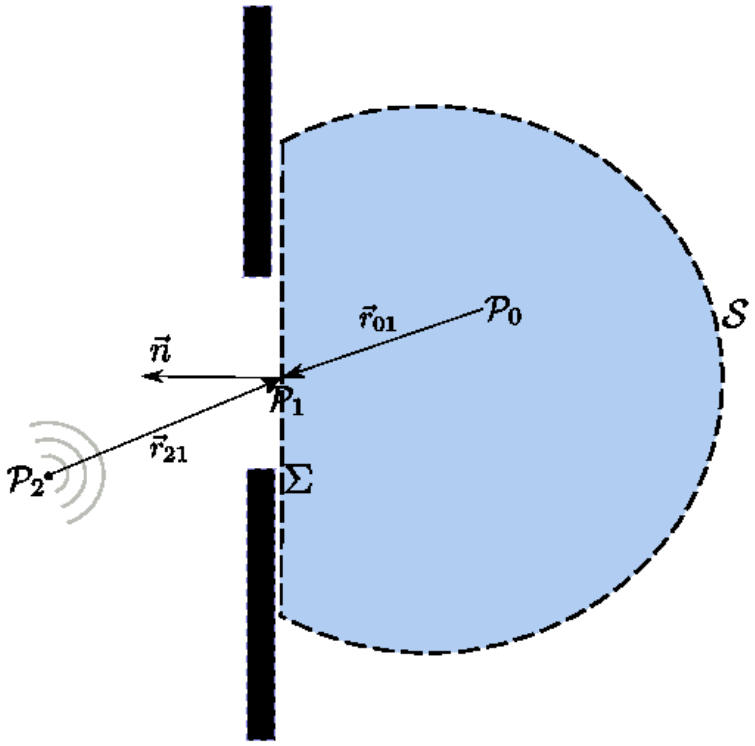


Figure 9: Diffraction by a planar screen illuminated by a single spherical wave (© Ariana-INRIA/I3S).

at an arbitrary point $\mathcal{P}_1$ that is a solution to Eq. (13) is

$$G(P_1) = \frac{\exp(jk_0 r_{o1})}{r_{01}}. \tag{14}$$

Here, $\mathbf{r}_{01}$ is the distance from the aperture Σ to the observation point P_0, and $k_0 = 2\pi/\lambda$ is the wave number in vacuum.

$$\frac{\partial}{\partial \mathbf{n}} G(P_1) = \cos(\mathbf{n}, \mathbf{r}_{01})(jk_0 - \frac{1}{r_{01}})\frac{\exp(jk_0 r_{01})}{r_{01}}, \tag{15}$$

where, $\mathbf{n}$ is the outward normal to Σ, and $\cos(\mathbf{n}, \mathbf{r}_{01})$ is the cosine of the angle between the normal $\mathbf{n}$ and $\mathbf{r}_{01}$.

Theorem 2.2. *The integral theorem of Helmholtz Eq.* (10) *and the Kirchoff Eq.* (14) *give the field at any point P_0 expressed in terms of the boundary values of the wave on any closed surface $\mathcal{S}$ surrounding that point. Accordingly*

$$U(P_0) = \frac{1}{4\pi} \oint_{\mathcal{S}} \frac{\exp(jk_0 r_{01})}{r_{01}} \frac{\partial}{\partial \mathbf{n}} U - U \frac{\partial}{\partial \mathbf{n}} \frac{\exp(jk_0 r_{01})}{r_{01}} d\mathcal{S}. \tag{16}$$

This represents the basic equation of scalar diffraction theory.

Fresnel-Kirchoff Diffraction Formula

Theorem 2.3. *If we assume that an aperture Σ is illuminated by a single spherical wave originating from $\mathcal{P}_2$ and that $r_{01} \gg \lambda$, $r_{21} \gg \lambda$ the disturbance at $\mathcal{P}_0$ with Kirchoff's boundary conditions is [16]*

$$U(P_0) = \frac{A}{j\lambda} \int_{\Sigma} \frac{\exp(jk_0(r_{21} + r_{01}))}{r_{21} r_{01}} \Big(\frac{\cos(\mathbf{n}, \mathbf{r}_{01}) - \cos(\mathbf{n}, \mathbf{r_{21}})}{2} \Big) d\mathcal{S}. \tag{17}$$

Proof. By theorem 2.2 and applying the Kirchoff's boundary conditions we get

$$U(P_0) = \frac{1}{4\pi} \int_{\Sigma} (G \frac{\partial}{\partial \mathbf{n}} U - U \frac{\partial}{\partial \mathbf{n}} G) d\mathcal{S}. \tag{18}$$

The fringing effects can be neglected if the dimension of the aperture is much larger than the wavelength λ of the light used. If $k_0 \gg 1/r_{01}$, then Eq. (15) can be written as

$$\frac{\partial}{\partial \mathbf{n}} G(P_1) \approx jk_0 \cos(\mathbf{n}, \mathbf{r_{01}}) \frac{\exp(jk_0 r_{01})}{r_{01}}. \tag{19}$$

Substituting Eq. (19) in Eq. (18), we get

$$U(P_0) = \frac{1}{4\pi} \int_{\Sigma} \frac{\exp(jk_0 r_{01})}{r_{01}} \Big(\frac{\partial}{\partial \mathbf{n}} U - jk_0 U \cos(\mathbf{n}, \mathbf{r_{01}}) \Big) d\mathcal{S}. \tag{20}$$

If we assume that the aperture is illuminated by a single spherical wave arising from P_2, then

$$U(P_1) = A \frac{\exp(jk_0 r_{21})}{r_{21}}. \tag{21}$$

Hence

$$U(P_0) = \frac{A}{j\lambda} \int_{\Sigma} \frac{\exp(jk_0(r_{21} + r_{01}))}{r_{21} r_{01}} \Big(\frac{\cos(\mathbf{n}, \mathbf{r}_{01}) - \cos(\mathbf{n}, \mathbf{r_{21}})}{2} \Big) d\mathcal{S}. \tag{22}$$

□

Remarks

(a) By the reciprocity theorem of Helmholtz, the effect of placing the point source at P_2 and observing at P_0 is equivalent to placing the point-source at P_0 and observing at P_2.

(b) The Fresnel-Kirchoff diffraction formula essentially confirms the Huygens principle. The field at P_0 arises from an infinite number of fictitious secondary point sources located within the aperture itself. The secondary sources here contain amplitudes and phases that are related to the illumination wavefront, and the angles of illumination and observation.

(c) The Fresnel-Kirchoff diffraction approximation is similar to the Rayleigh-Sommerfield theory for small diffraction angles.

Debye approximation From the Kirchoff-Fresnel formulation, the Debye integral approximation for a circular aperture can be obtained as [15]

$$U(P_0) = \frac{-j}{\lambda} \int_0^{2\pi} \int_0^{\alpha} A(\theta_i) \exp(-jk_i \rho \sin\theta_i \cos(\phi - \psi)) \exp(jk_i z \cos\theta_i) \sin\theta_i d\theta_i d\phi, \tag{23}$$

where α is the semi-aperture angle of the objective, A is the *apodization function*, $k_i = 2\pi n_i/\lambda$, and λ/n_i is the wavelength in the medium of refractive index n_i. For a lens with a uniform aperture, the apodization function is radially symmetrical with respect to the optic axis and can be represented by $A(\theta_i)$. The intensity projected from an isotropically illuminating point source such as a flurophore, on a (flat) pupil plane is bound to be energy conservation constraint. Therefore the amplitude as a function of θ_i in the pupil plane should vary as $(\cos\theta_i)^{-1/2}$ and the energy as $(\cos\theta_i)^{-1}$ [6]. Thus, $A(\theta_i)$ for the detection is given by

$$A(\theta_i) = (\cos\theta_i)^{-\frac{1}{2}}, \tag{24}$$

and for the illumination as

$$A(\theta_i) = (\cos\theta_i)^{\frac{1}{2}}. \tag{25}$$

Eqs. (24) and (25) are known as the "*sine condition*" and they guarantee that a small region of the object plane in the neighborhood of the optic axis is imaged sharply by a family of rays which can have any angular divergence. This constitutes an aplanatic imaging system and it exhibits 2D transverse shift invariance.

Remarks The Debye approximation holds good only if

(a) $r_{21} \gg a$, (a is the radius of the aperture)

(b) the spherical wavelets from the aperture Σ are approximated by plane wavelets,

(c) $\cos(\mathbf{n}, \mathbf{r}_{01}) \approx -1$,

(d) the numerical aperture (NA) is very large and the Fresnel number is high.

Stokseth approximation Stokseth [14] also arrived at the above approximation, in Eq. (23), by extending the work of Hopkins [17]. Hopkins essentially worked out an approximation for small amounts of defocusing. While Stokseth's is a simple analytical approximation for large defocusings. The PSF is defined here as the irradiance distribution in the image plane of a point source in the object plane. As the OTF for incoherent illumination and the PSF are related by Fourier transforms, we say

$$h(\mathbf{x}) = \mathcal{F}^{-1}[OTF(\mathbf{k})]. \tag{26}$$

The complex amplitude PSF can hence be written as

$$h_A(\mathbf{x}) = \int_{\mathbf{k}} OTF_A(\mathbf{k}) \exp(j\mathbf{k}\cdot\mathbf{x}) d\mathbf{k}, \tag{27}$$

where $\mathbf{x}$ and $\mathbf{k}$ are the 3D coordinates in the image and the Fourier space. By making the axial Fourier space co-ordinate k_z a function of lateral co-ordinates, $k_z = (k_i^2 - (k_x^2 + k_y^2))^{1/2}$, the 3D Fourier transform is reduced to

$$h_A(x,y,z) = \int_{k_x}\int_{k_y} P(k_x,k_y,z) \exp(j(k_x x + k_y y)) dk_y dk_x, \tag{28}$$

where $P(\cdot,\cdot,\cdot)$ describes the overall complex field distribution in the pupil of an unaberrated objective lens. The pupil function is a description of the magnitude and phase of the wavefront that a point source produces at the exit pupil of the imaging system. In simple terms the above expression states that the field distribution at a point (x,y,z) in the image space can be obtained by applying Fourier transform on the overall pupil function. For a microscope suffering from defocus, the pupil function can be written as

$$P(k_x,k_y,z) = \begin{cases} A(\theta_i)\exp(jk_0\varphi(\theta_i,\theta_s,z)), & \text{if } \frac{\sqrt{k_x^2+k_y^2}}{k_i} < \frac{NA}{n_i} \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

where $\varphi(\theta_i,\theta_s,z)$ is the optical phase difference between the wavefront emerging from the exit pupil and the reference sphere measured along the extreme ray, and $\theta_i = \sin^{-1}(k_x^2+k_y^2)^{1/2}/k_i$ and $\theta_s = \sin^{-1}(k_x^2+k_y^2)^{1/2}/k_s$ (see fig. 12). θ_s and θ_i are related by Snell's law as

$$n_i \sin\theta_i = n_s \sin\theta_s. \tag{30}$$

In fact, $\varphi(\theta_s,\theta_s,z)$ is a sum of two terms, the defocus term $\varphi_d(\theta_s,z)$ and the term due to aberrations $\varphi_a(\theta_i,\theta_s)$. For spherical aberration free imaging conditions, it is to be noted that $\theta_i \approx \theta_s$, and hence $\varphi_a(\theta_i,\theta_s) \approx 0$. This will be discussed in subsection 2.2.1.

If f is the distance between the exit pupil and the geometrical image point, then Stokseth obtained the exact expression for this defect of focus as

$$\varphi_d(\theta_s,z) = -f - z\cos\theta_s + (f^2 + 2fz + z^2\cos^2\theta_s)^{\frac{1}{2}}. \tag{31}$$

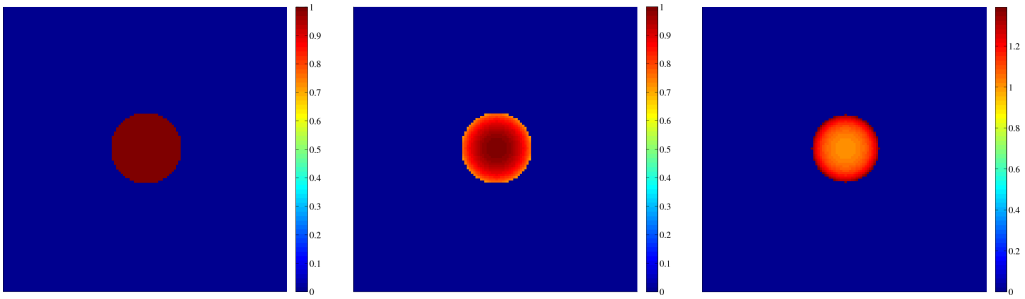


Figure 10: The amplitude of the pupil function of a CLSM for the (a) ideal, (b) illumination, and (c) emission cases (© Ariana-INRIA/CNRS).

A small angle approximation can be obtained by making a series expansion of $\cos\theta_s$ and $\cos^2\theta_s$, omitting terms of θ_s with higher order than two

$$\varphi_d(\theta_s, z) = \frac{1}{2} z\theta_s^2 \Big(1 - \frac{z}{f+z}\Big), \theta_s \ll 1. \tag{32}$$

For small defocusings and small angles $z^2\cos^2\theta_s \approx z^2$, and hence Eq. (31) becomes

$$\varphi_d(\theta_s, z) = z(1 - \cos\theta_s) \tag{33}$$

We rewrite in the spherical coordinates as

$$\mathbf{k}\cdot\mathbf{x} = k_i \sin\theta_i (x\cos\phi + y\sin\phi), \tag{34}$$

and

$$d^2\mathbf{k} = k_i^2 \sin\theta_i \cos\theta_i d\theta_i d\phi. \tag{35}$$

By using Eqs. (29), (33), (34), (35) into Eq. (28), we get:

$$h_A(x,y,z) = k_i^2 \int_0^{\alpha}\int_0^{2\pi} (\cos\theta_i)^{-\frac{1}{2}} \exp(jk_i \sin\theta_i (x\cos\phi + y\sin\phi)) \cdot \exp(jk_s z(1-\cos\theta_s)) \sin\theta_i \cos\theta_i d\theta_i d\phi. \tag{36}$$

If $x = \rho\cos\psi$ and $y = \rho\sin\psi$, then $\rho = (x^2+y^2)^{1/2}$, and

$$h_A(x,y,z) = k_i^2 \int_0^{\alpha}\int_0^{2\pi} (\cos\theta_i)^{\frac{1}{2}} \exp(jk_i\rho\sin\theta_i\cos(\phi-\psi)) \exp(jk_s z(1-\cos\theta_s)) \sin\theta_i d\theta_i d\phi. \tag{37}$$

Remarks There are subtle differences between Eqs.(23) and (37), namely:

(a) The extra term $\exp(jkz)$ and $\exp(j\pi/2)$ is canceled when the complex amplitude is squared to get the PSF.

(b) The final intensities are normalized so that they sum to unity.

2.2 Theory of Aberrations

Under ideal conditions, a high NA objective lens transforms the planar wavefront incident on it to a spherical wavefront at the focal region. However, under practical situations, the refracted wavefront so produced has to go through several optical elements and the specimen. Due to this reason the emerging wavefront is rarely spherical in nature. Aberrant wavefront means that the resulting observed images will be distorted as well. While there are many aberrations that exist for the microscope, we restrict our analysis to spherical aberration (SA) as this is the dominant and the most observable form.

2.2.1 Spherical Aberrations

Spherical aberration (SA) is an optical effect occurring when the oblique rays entering a lens are focused in a different location than the central rays. The mismatch between the refractive index of the objective lens immersion medium and the specimen embedding medium causes SA in fluorescence microscope. When light crosses the boundary between materials with different refractive indices, it bends across the boundary surface differently depending on the angle of incidence (light refraction); the oblique rays are bent more than the central rays. If the mismatch is large, e.g. when going from oil lens immersion medium into a watery specimen embedding medium, SA causes the PSF to become asymmetric at depths of even a few microns. Also, the amount of light collected from a point source decreases

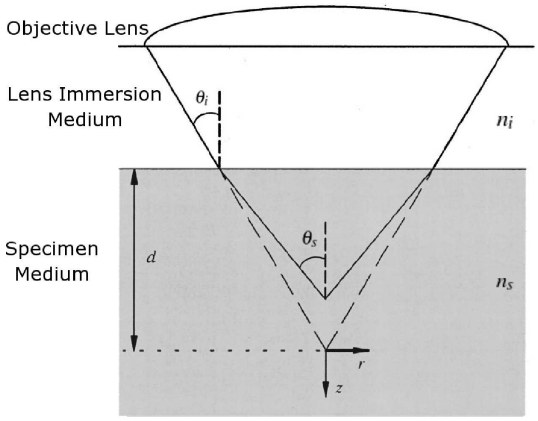


Figure 11: Schematic describing the focusing of light when traveling between medium of different refractive indices (© Ariana-INRIA/I3S).

with depth because of an axial broadening of the PSF. When measured experimentally using fluorophores [18], it was found that the PSF changes from a fairly symmetric axial shape to an asymmetric shape. It is important to remember the following features of a depth varying PSF (DVPSF)

- in the absence of other aberrations, the PSF remains radially symmetrical,
- its peak intensity decreases with increase in the depth of focus d,
- there is an increase in the width of the PSF in the axial but not the lateral direction with increase in the depth of focus d,
- the total amount of light collected from a given object does not change with depth. This is because the decrease in peak amplitude is coincident with an increase in the profile width such that the total collected light remains essentially constant.

We consider the situation where the objective lens focuses through an interface between media with different refractive indices as shown in Fig. 11. The distance d is the nominal focusing depth, (i.e. the depth of focus in a matched medium) and the focus is made from a medium of refractive index μ_i to a medium of refractive index μ_s. Practical imaging systems are rarely aberration-free. With aberrations, the wavefront emerging from the lens is not a spherical surface. Under such conditions, the imaging system also loses its property of shift-invariance.

2.2.2 The Phase Factor

Consider a point source located at a depth d below the cover slip in a mounting medium of index n_s, observed with an objective lens designed for an immersion medium with a refractive index n_i (see Fig. 12). In Eq. (37), we had ignored the aberrations by assuming that $n_i \approx n_s$ and hence $\theta_i \approx \theta_s$. When this is not the case, we have to calculate the aberration function $\varphi_a(\theta_i, \theta_s)$, due to the mismatch of n_s and n_i. The phase change is determined by the difference between the optical path length traveled to the pupil by a ray that leaves the source at an angle θ_s relative to the optic axis and is refracted to the angle θ_i upon leaving the mounting medium, and the optical path length that a ray with angle θ_i would have traveled if the mounting medium index were the ideal index n_i. In their paper, Gibson and Lanni [19] mention that there are 8 parameters (out of a total of 18) that may vary from their design conditions as recommended by the microscope manufacturer. However, when a microscope is properly calibrated, there are only 3 parameters that essentially vary under experimental conditions. These are

1. the specimen refractive index n_s,
2. the immersion medium refractive index n_i of the objective lens, and
3. the depth d under the coverslip where lies the plane of focus.

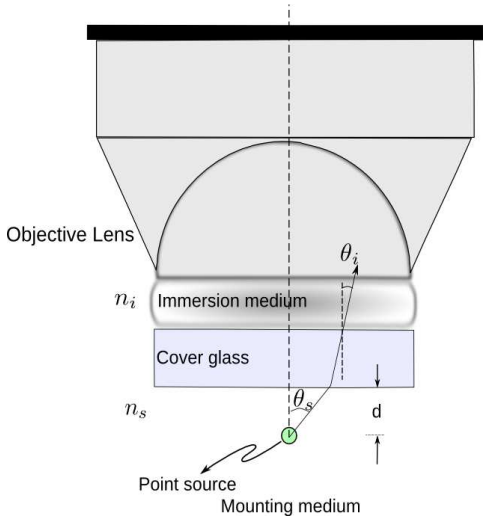


Figure 12: Spherical aberration due to refractive index mismatch $n_s \neq n_i$ (© Ariana-INRIA/I3S).

The refractive index of the immersion medium is sensitive to changes in temperature especially if an oil immersion lens is used. An error in this parameter significantly affects the numerically computed PSF.

Using simple geometrical optics, we can show that this shift in the optical path can be analytically expressed as

$$\varphi_a(\theta_i, \theta_s; d, n_i, n_s) = d(n_s \cos\theta_s - n_i \cos\theta_i), \tag{38}$$

where d is the focusing at a nominal depth into the specimen of refractive index n_s from a medium of refractive index n_i. The above phase term relies on the assumption that the error due to mismatch in the refractive indices between the cover glass n_g and the objective lens has either been compensated by the correction collar or is minimal. If the cover glass is used with an objective lens that is significantly different than its design specification, an additional phase term should be included, and d replaced by the thickness of the coverslip.

2.3 Approximating the Point-Spread Function

2.3.1 Nonlinear Phase Approximation

The phase term in Eq. (38) can be rewritten as

$$\begin{aligned} \varphi_a(\theta_i, \theta_s; d, n_i, n_s) &= d \sec\theta_i (n_s \cos\theta_s \cos\theta_i - n_i \cos^2\theta_i), \\ &= d \sec\theta_i (n_s \cos\theta_s \cos\theta_i - n_i (1 - \sin^2\theta_i)), \end{aligned} \tag{39}$$

as $(\sin^2\theta_i + \cos^2\theta_i) = 1$. But since $n_s \sin\theta_s = n_i \sin\theta_i$ from the Snell's law and $\cos(\theta_s - \theta_i) = (\cos\theta_s \cos\theta_i + \sin\theta_s \sin\theta_i)$, the phase is

$$\begin{aligned}\varphi_a(\theta_i, \theta_s; d, n_i, n_s) &= d \sec\theta_i (n_s \cos\theta_s \cos\theta_i - n_i + n_s \sin\theta_s \sin\theta_i), \\ &= d \sec\theta_i (n_s \cos\theta_s \cos\theta_i - n_i + n_s \sin\theta_s \sin\theta_i), \\ &= d \sec\theta_i (n_s \cos(\theta_s - \theta_i) - n_i) \end{aligned} \tag{40}$$

If we assume that the aberration is not very large, *i.e.* $\theta_s \approx \theta_i$, so that

$$\varphi_a(\theta_i, \theta_s; d, n_i, n_s) \approx d \sec\theta_i (n_s - n_i). \tag{41}$$

2.3.2 Linear Phase Approximation

The phase term in Eq. (38) can be approximated by using a set of circular basis functions called Zernike polynomials [20]. These polynomials form a complete orthogonal set on the unit disk. Since there is no azimuthal variation, it is sufficient to consider Zernike circle polynomials of order n and zero kind $(Z_n^0(\rho; d, n_i))$. Here $\rho\,(0 \le \rho \le 1)$ is the normalized radial coordinate.

$$\varphi_a(\rho; d, n_i) = d\,\mathrm{NA}\Big(\sum_{n=0}^{4} c_n Z_n^0(\rho; d, n_i)\Big) \tag{42}$$

where the expansion coefficients (c_n) need to be calculated or determined. Since we are interested only in defocus and the first order SA, $\varphi_a(\rho; d, n_i, n_s)$ is approximated to only the Zernike polynomials of order up to 4.

2.3.3 Apodization Function Approximation

For the gradient calculations to be discussed in Appendix C, we make an approximation on the apodization function $A(\theta_i)$ of Eq. (29). We simply assume the apodization function for excitation and emission are approximately same and equal to one. That is

$$A(\theta_i) \approx 1. \tag{43}$$

2.4 Analytical Expression

In this section we summarize the theory that has been discussed earlier for obtaining the analytical PSF expression. We proceed first with the following definitions,

Definition Bessel function of first kind and integer order n

$$\forall\ x \in \mathbb{R},\ n \in \mathbb{N},\ J_n(x) = \frac{1}{\pi} \int_0^{\pi} \cos(n\theta - x\sin\theta) d\theta \tag{44}$$

Corollary 2.4. *Bessel function of first kind with order zero*

$$\forall\ x \in \mathbb{R},\ J_0(x) = \frac{1}{\pi}\int_0^{\pi} \cos(x\sin\theta)d\theta \tag{45}$$

Thus for the nonparaxial case (i.e. NA>0.7), Eq. (28) can be modified to now include the additional phase term as

$$h_A(\mathbf{x};\lambda) \propto \int_0^{\alpha} A(\theta_i)\sin\theta_i J_0(k_i\rho\sin\theta_i)\exp(jk_0\varphi_d(\theta_s,z;n_i))\exp(jk_0\varphi_a(\theta_i,\theta_s;d,n_i,n_s))d\theta_i. \tag{46}$$

and

$$h_A(\mathbf{x};\lambda) = 2\pi k_i^2\exp(jk_sz)\int_0^{\alpha}(\cos\theta_i)^{\frac{1}{2}}\sin\theta_i J_0(k_i\rho\sin\theta_i)\cdot \tag{47}$$
$$\exp(jk_0\varphi_a(\theta_i,\theta_s;d,n_i,n_s))\exp(-jk_sz\cos\theta_s)d\theta_i.$$

By using a normalized radial co-ordinate $t = \sin\theta_i/\sin\alpha$, the field at P_0 becomes

$$h_A(\mathbf{x};\lambda) = 2\pi k_i^2\exp(jk_sz)\int_0^{1}(1-t^2\sin^2\alpha)^{-\frac{1}{2}}t\sin^2\alpha J_0(kt\rho\sin\alpha)\cdot \tag{48}$$
$$\exp(jk_0\varphi_a(t,\theta_s;d,n_i,n_s))\exp(-jk_sz(1-t^2\sin^2\alpha)^{\frac{1}{2}})dt.$$

For small angles (i.e. NA≤ 0.7), $\sin\theta_i \approx \theta_i$ and hence

$$Eq.(48) = 2\pi k_i^2\sin^2\alpha\int_0^{1} J_0(kt\rho\sin\alpha)\exp(jk_0\varphi(t,\theta_s;d,n_i,n_s))tdt, \tag{49}$$

where

$$\varphi(t,\theta_s;d,n_i,n_s) = \varphi_a(t,\alpha,\theta_s;d,n_o,n_s) + \varphi_d(t,z,\alpha), \tag{50}$$

and

$$\varphi_d(t,z,\alpha) = \frac{k_sz\sin^2\alpha}{2k_0}t^2 \tag{51}$$

is the quadratic defocus term.

The Fourier transform of a circularly symmetrical function $g(\rho)$ written as $\mathcal{B}(g(\rho))$ can be given as

$$\mathcal{B}(g(\rho)) = G_k = 2\pi\int_0^{\infty} rg(r)J_0(2\pi rk)dr \tag{52}$$

where, J_0 is the bessel function of the first kind of order zero. For a circularly symmetrical function $f(\rho,\theta) \equiv f(\rho)$, the *Fourier-Bessel transform* or *Hankel transform of zero order* is given as:

$$\mathcal{B}[f(\rho)] = 2\pi \int_0^\infty \rho f(\rho) J_0(2\pi\rho k) d\rho. \tag{53}$$

Thus the Fourier transform of a circularly symmetric function is itself circularly symmetrical.

Definition Circle function

$$circ(\rho) = \begin{cases} 1, & \rho < 1 \\ \frac{1}{2}, & \rho = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{54}$$

The Fourier-Bessel transform of the circle function can be written as

$$\mathcal{B}[circ(\rho)] = 2\pi \int_0^1 \rho J_0(2\pi\rho k) d\rho \tag{55}$$

Using a change of variable $\rho' = 2\pi\rho k$ and the identity $\int_0^x \xi J_0(\xi) d\xi = x J_1(x)$,

$$\mathcal{B}[circ(\rho)] = \frac{1}{2\pi k^2} \int_0^{2\pi k} \rho' J_0(\rho') dr' = \frac{1}{k} J_1(2\pi k) \tag{56}$$

where, $J_1(\cdot)$ is a Bessel function of the first kind and order one. By using the notation of Fourier-Bessel transform or Hankel transform of zero-order, Eq. (49) is simplified as

$$h_A(\mathbf{x};\lambda) = k_i^2 \sin^2\alpha \, \mathcal{B}[\exp(jk_0\varphi(t,\theta_s;d,n_i,n_s))]. \tag{57}$$

Summarizing the results obtained earlier, the amplitude distribution can be either written as

$$h_A(\mathbf{x}) \propto \int_0^\alpha A(\theta_i)\sin(\theta_i) J_0(k_i\rho\sin\theta_i) \exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s)) d\theta_i, \tag{58}$$

which is a simplification of Eq. (37) or Eq. (49) as (by ignoring apodization)

$$h_A(x,y,z;\lambda) \propto \int_0^1 J_0(k_i\rho\sin\alpha t) \exp(jk_0\varphi_d) \exp(jk_i\varphi_a) t dt, \tag{59}$$

where J_0 is the Bessel function of the first kind of order zero, $\varphi_d = (z\sin^2\alpha t^2)/2$, $\varphi_i = d(n_s\cos\theta_s - n_i\cos\theta_i)$, $n_i\sin\theta_i = n_s\sin\theta_s$ and $n_i\sin\alpha = \text{NA}$. $u = (2\pi n_o/\lambda)\rho\sin\alpha$ and $v = (2\pi n_s/\lambda) z\sin\alpha$ are the optical coordinates.

For a single photon (1-p) excitation, when the fluorophore returns to the ground state the emitted wavelength is longer than the excitation wavelength (*Stoke's shift*). From the Helmholtz reciprocity theorem, the PSF is written as a product of the excitation distribution and the emission distribution as

$$h_{Th} \propto |h_A(\mathbf{x};\lambda_{ex})|^2 \cdot \int\limits_{x_1^2+y_1^2 \leq \frac{D^2}{4}} |h_A(x-x_1, y-y_1, z;\lambda_{em})|^2 dx_1 dy_1, \tag{60}$$

where λ_{ex} and λ_{em} are respectively the excitation and emission wavelengths, D is the back projected diameter of the circular pinhole. Fig. 13 shows a numerically computed PSF by using the Eqs. (33), (38), (50) and (58) in Eq. (60). The microscope uses an excitation beam of wavelength $\lambda_{ex} = 488$nm, and the emission peak has a wavelength of $\lambda_{em} = 520$nm. The objective lens has a NA of 1.3, 40X magnification and is an immersion oil type refractive index $n_i = 1.518$.

In incoherent imaging, the distribution of intensity in the image plane is found by integrating the intensity distributions in the diffraction images associated with each point in the object. Thus if $o(\mathbf{x}')$ is the intensity at $\mathbf{x}'$ in the object plane, the intensity at the point $\mathbf{x}$ in the image is obtained as (ignoring the influence of noise):

$$i(\mathbf{x}) = \int\limits_{\mathbf{x}\in\Omega_s} o(\mathbf{x}')h(\mathbf{x}-\mathbf{x}')d\mathbf{x}'. \tag{61}$$

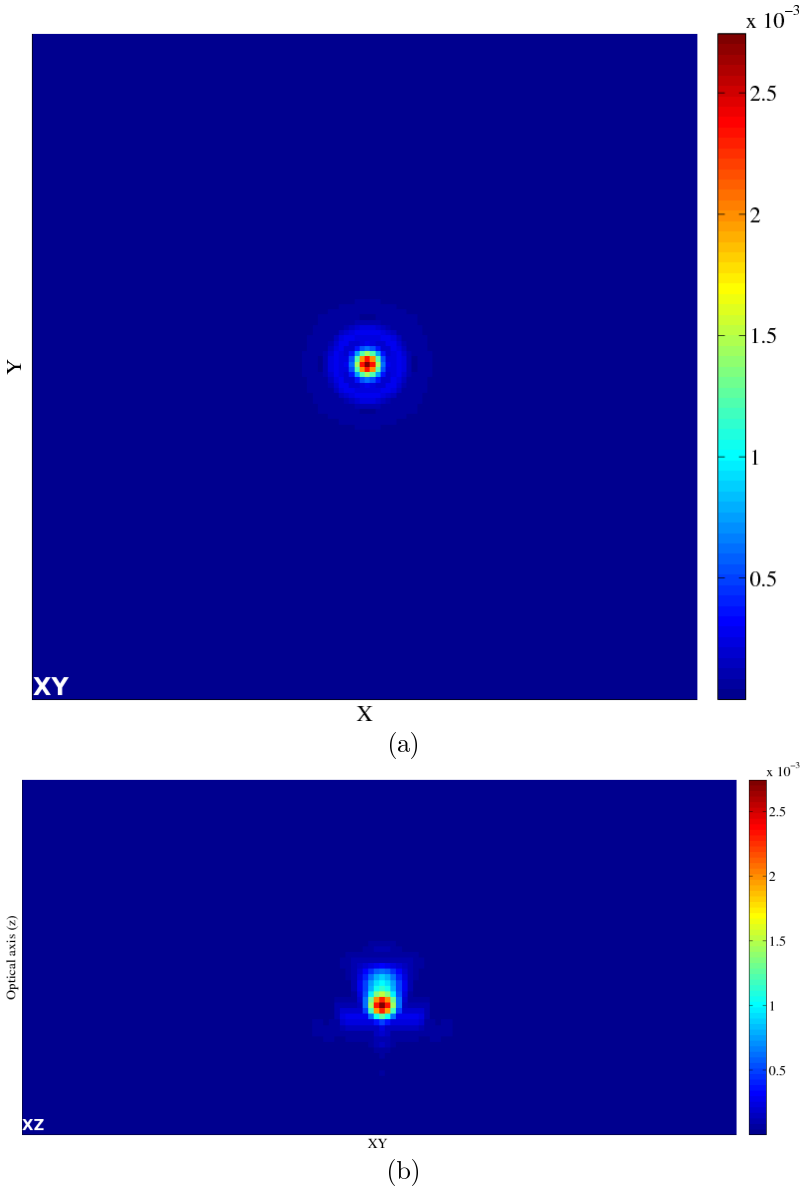


Figure 13: Maximum intensity projection (MIP) of a numerically computed spherically aberrated confocal PSF along the (a) optic axis giving the radial plane, (b) lateral axis giving the axial plane (© Ariana-INRIA/CNRS).

3 Literature Review on Blind Deconvolution

In this section we review some existing methods on blind deconvolution applied to microscopy. It is difficult to make an exhaustive review of all the literature available in this field. So, we have chosen only those that resemble the problem we are handling or methods that have raised considerable interests on this subject. We have skipped the survey on deconvolution algorithms as this exercise was carried out earlier in [21, 22] and more recently in our research report [8] with different object regularization.

3.1 Maximum Likelihood Approach

If we assume that the observed image $i(\mathbf{x})$ is a realization of an independent Poisson process at each voxel, then the likelihood can be written as

$$\Pr(i|o,h) = \prod_{\mathbf{x}\in\Omega_s} \frac{[(h*o)+b](\mathbf{x})^{i(\mathbf{x})} e^{-[(h*o)+b](\mathbf{x})}}{i(\mathbf{x})!}, \tag{62}$$

where, the mean of the Poisson process is given by $[(h*o)+b](\mathbf{x})$. Conventional blind deconvolution algorithms estimate the object and the PSF directly from the above equation. That is

$$(\hat{o},\hat{h}) = \arg\max_{(o,h)}\{\Pr(i|o,h)\} \tag{63}$$

A maximum likelihood (ML) approach can be used to recover both the object and PSF as

$$\begin{aligned} \hat{o}_{\mathrm{ML}} &= \arg\max_{o}\{\Pr(i|o,h)\} \\ &= \arg\min_{o}\{-\log(\Pr(i|o,h))\}. \end{aligned} \tag{64}$$

$$\begin{aligned} \hat{h}_{\mathrm{ML}} &= \arg\max_{h}\{\Pr(i|o,h)\} \\ &= \arg\min_{h}\{-\log(\Pr(i|o,h))\}. \end{aligned} \tag{65}$$

An explicit iterative multiplicative algorithm based on Maximum Likelihood Expectation Maximization (MLEM) [9] formalism can be derived from the above expression as

$$\hat{o}^{n+1}_{\mathrm{ML}}(\mathbf{x}) = \hat{o}^{n}_{\mathrm{ML}}(\mathbf{x}) \cdot \left(\frac{i(\mathbf{x})}{(\hat{o}^{n}_{\mathrm{ML}} * h)(\mathbf{x})} * h(-\mathbf{x}) \right)^{q_o}, \forall \mathbf{x} \in \Omega_s. \tag{66}$$

The method involves alternating between object minimization above and the PSF minimization

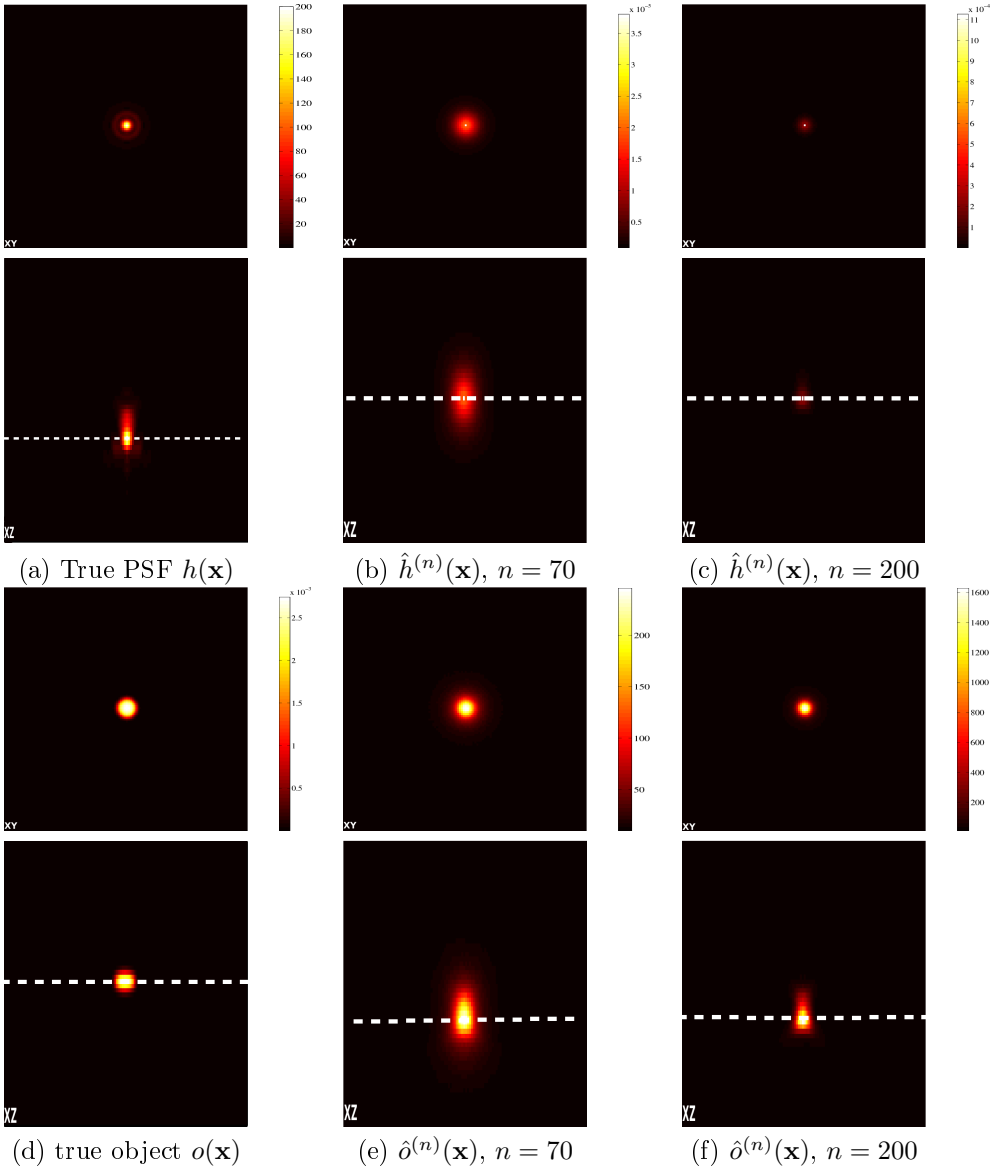
$$\hat{h}^{n+1}_{\mathrm{ML}}(\mathbf{x}) = \hat{h}^{n}_{\mathrm{ML}}(\mathbf{x}) \cdot \left(\frac{i(\mathbf{x})}{(o * \hat{h}^{n}_{\mathrm{ML}})(\mathbf{x})} * o(-\mathbf{x}) \right)^{q_h}, \forall \mathbf{x} \in \Omega_s, \tag{67}$$

where n is the index of iteration of the algorithm. For those unfamiliar with MLEM, its derivation from the Bayes's theorem and as formalized by Richardson [23] is summarized in Appendix A. The parameters $q_o, q_h \in [1, 1.5]$ controls the convergence of the two iterations. When q_o and q_h is unity, then we arrive at the classical MLEM algorithm and when they are more than 3, they converge (and sometimes diverge) rapidly.

In Fig. 14, we show the results of applying Eqs. (66) and (67) to the observation of Fig.5(c). As this algorithm has to be manually terminated, we show the results for two iterations with $n = 70$ and $n = 200$. Fig. 14(a) and (d) gives the true aberrated PSF and the object used for simulating the observation. Since the algorithm does not have any information about the object or the PSF (other than positivity and flux conservation), the estimate of the object $\hat{o}^{(n)}(\mathbf{x})$ in Fig. 14(e) and (f), resembles closely the true PSF in Fig. 14(a), in shape and in the position of the COG. Similarly, the estimate of the PSF $\hat{h}^{(n)}(\mathbf{x})$ in Fig. 14(b) and (c) resembles more the imaged axially centered microsphere in Fig. 14(d). Thus, having no prior knowledge on the object or the PSF can make it difficult to separate or distinguish them from each other. From this simple example, we clearly observe that in the MLEM estimation, the characteristics of the PSF is absorbed by the object and vice versa. Also, when the iterations are allowed to continue, progressively the results start deteriorating as can be seen from Fig. 14(c) after 200 iterations of the accelerated blind deconvolution. We perform another experiment, where the object is axially shifted from its center by two planes as shown in Fig. 15(a). Fig. 15(c) shows the same true PSF as in Fig. 14(a). Fig. 15(b) and (d) show $\hat{o}^n(\mathbf{x})$ and $\hat{h}^n(\mathbf{x})$ after $n = 70$ iterations. The results of this experiment leads us to conclude that the true axial position of the object cannot also be recovered using this method of MLEM blind deconvolution.

3.2 Penalized Maximum Likelihood Approach

In the classical MLEM, the stopping criteria is usually the number of iterations. This is a constraint introduced in the algorithm that prevents early divergence of the results. The earliest blind deconvolution approach that was applied to incoherent quantum-limited imaging and application to fluorescent microscopes was by Holmes in [24]. At each iteration of the MLEM for the PSF, a unit summation constraint is enforced on the the PSF. In addition, the energy of the PSF is constrained to lie within an hourglass-shaped region by using a Gerchberg-Saxton [25] type algorithm. This prevents portion of the background fluorescence $b(\mathbf{x})$ from being erroneously considered as having originated from the PSF. Finally, a bandlimiting and positivity criterion is introduced. This form of grouped coordinate descent is typically referred to as an iterative blind deconvolution (IBD). Biggs and Andrews [26] provided an alternative approach for accelerating the IBD. Since the object estimate converges faster than the PSF estimates, a modification was proposed wherein the acceleration was achieved by performing several iterations for the PSF after each object estimation. The number of cycles of PSF iterations to apply after each iteration of the object estimate was experimentally chosen. Given an observation data, the estimation of the parameters $\{q_o, q_h\}$ or the number of iterations in Biggs method [26] remains still an open problem. Although



(a) True PSF $h(\mathbf{x})$ (b) $\hat{h}^{(n)}(\mathbf{x})$, $n = 70$ (c) $\hat{h}^{(n)}(\mathbf{x})$, $n = 200$

(d) true object $o(\mathbf{x})$ (e) $\hat{o}^{(n)}(\mathbf{x})$, $n = 70$ (f) $\hat{o}^{(n)}(\mathbf{x})$, $n = 200$

Figure 14: Blind deconvolution results based on the MLEM algorithm after 70 and 200 iterations (© Ariana-INRIA/I3S).

this algorithm works well, it suffers from some major drawbacks. Secondly, it requires infor-

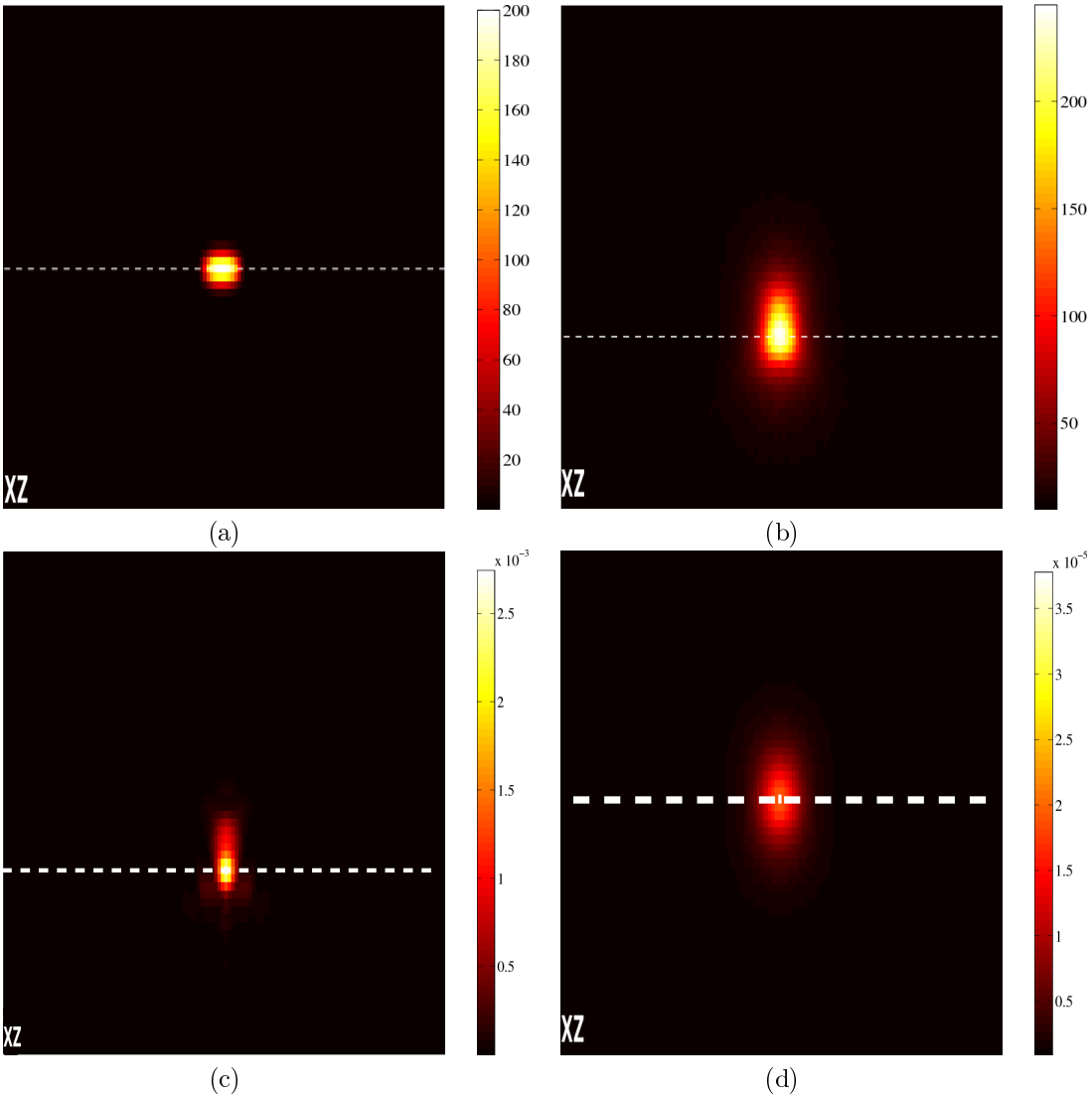


Figure 15: (a) Simulated true object $o(\mathbf{x})$ that has been shifted along the optic axis of the volume, (b) the estimated object $\hat{o}(\mathbf{x})$ using the MLEM blind deconvolution, (c) the true PSF $h(\mathbf{x})$ used for simulation, (d) the estimated PSF $\hat{h}(\mathbf{x})$ using the MLEM blind deconvolution. Here the number of iterations $n = 70$ (© Ariana-INRIA/I3S).

mation about the experimental setup to place constraints on the spatial and spectral extent of the PSF. In many cases this might not be available or might change with experimental conditions. Thirdly, in its original form, it cannot be applied when there are aberrations (for fluorescent microscope see Subsection 2.2) in the imaging process.

Incorporation of prior knowledge in the reconstruction process through a prior function makes the problem well-posed. The emerging research has been in the search for a good prior that models the object and the PSF [27]. Towards this end there have been several prior representations of the object based on either ℓ_1 [28, 10], ℓ_2, ℓ_1-ℓ_2 [29] norms, or entropy [30], wavelets [31], sparsity [32] and median root priors (MRP) [33, 34]. A very instructive study on this kind of regularization, especially applied to fluorescence microscopy, was carried out in [30]. Both Gaussian and Poisson models were considered for the noise in combination with Tikhonov, entropy, Good's roughness regularization and also with no regularization (MLE with uniform prior). It was shown that the Good's roughness was the best choice for the regularization between the priors considered, *viz.* ℓ_2-type and entropy. The Gaussian prior is widely used in fluorescence imaging because of its well-known properties of preventing noise amplifications and convergence. It can also be seen as the Tikhonov functional in the traditional regularization problem. It was also mentioned that the entropy distribution is optimal for non-negative functions in the absence of further prior knowledge on the function. However, for blind deconvolution applications a ℓ_2 kind of regularization on the object does not make the separation of the PSF and the object easy as the solution space is similar. It is one of the primary reasons why we had recommended using an ℓ_1 regularization in the form of Total Variational (TV) functional on the object [10]. Besides this, the TV regularization preserves the edge and prevents over smoothing of the solutions as in the case of regularization utilizing ℓ_2 norm. It has also been observed that natural image statistics rarely follow a ℓ_2 or Gaussian priors. In fact the sparsity prior of [32] is based on the fact that the gradients of natural images have heavy tailed distributions rather than Gaussian. As for the PSF, in confocal microscopy, most algorithms just impose bandlimiting and positivity constraints or ℓ_2 norms. For the former, the underlying assumption is that the PSF is uniformly distributed with certain restrictions. Other works involve spatially approximating the PSF by parametric models [35, 11] or by linearizing the phase term of the pupil function of the PSF [36]. Whatever be the prior knowledge, the important point that should be noted for application to blind deconvolution is that the prior terms should enable separation of the object and the PSF solution spaces.

3.3 Other Approaches

Most iterative blind deconvolution algorithms are based on the Bayesian inference and hence draw their inspiration from the above basic form of the IBD. The only distinction being the different constraints they impose on the PSF and the object. Another popular method is the *nonnegativity and support constraints recursive image filtering algorithm* or NAS-RIF [37, 38]. Although this algorithm is known to have good convergence properties, it cannot be used for deconvolving fluorescent microscope images as the PSF does not have an exact inverse. For a review of other existing methodologies in the field of blind deconvolution,

the interested reader may refer to the following articles [39, 40]. It was mentioned in [9] that nonlinear image deconvolution algorithms based on probabilistic techniques, outperform conventional linear techniques. They offer several advantages over classical deterministic techniques like less noise amplification and restoration beyond the diffraction limits. One of the fundamental disadvantage of deterministic techniques is their inability to restore the information beyond a certain cut off frequency. An interesting general survey on linear and nonlinear algorithms can be found in [41]. Well known linear and nonlinear image restoration methods have been studied in depth and it can be observed that they are based on the assumption that the noise is either a Poisson or a Gaussian process. Linear algorithms were shown to be related to the nonlinear methods through the partial derivative, with respect to the object, of a Poisson or a Gaussian likelihood function. Among the non-iterative blind deconvolution algorithms, we found APEX [21, 42] to be fastest in terms of computational time and applicability to a certain class of PSFs (low-exponent, Lévy stable probability density functions) with Gaussian noise. Although the defocus blurs fall under the above category, they are not applicable for fluorescent microscopes, as the PSFs are specimen dependent and rarely fall under the above category.

In section 5, we will see that the existing methods of deconvolution (blind or nonblind) fail in the presence of imaging system aberrations. A major problem when imaging at depth within a biological sample in confocal or nonlinear microscopy is the introduction of sample induced aberrations. We will see in Section 2 the importance of using the phase of the emerging refracted wavefront in restoring the observed objects. Some prior work related to this subject is on wavefront reconstruction using adaptive optics (AO) [43, 44, 45, 46, 47]. The above methods were based on physically adding a special conjugate element to compensate for the aberrations caused by the specimen or the optics. This can also be achieved computationally as was shown in [48, 49]. A literature review of the recent and emerging trends in this field was done by Booth in [50].

The algorithm presented here on PSF retrieval through phase retrieval is inspired from Hanser's work for WFM [6], but within a Bayesian framework.

4 Point-Spread Function Estimation

As explained in Section 2, the non-aberrated PSF model in Eq. (37) is valid only when imaging thin samples (depth $< 5\mu m$). However, when imaging deeper into biological tissues, there is a dramatic reduction in both the signal level and the resolution. While most systems are built to be diffraction-limited, it is not possible to ensure that they maintain this performance throughout. It was shown in [2] that when focusing into samples immersed in water by using an oil immersion objective, the signal intensity level falls to 40% at 5 mm and to 10% at 15 mm beneath the cover glass! We propose a modification to the PSF expression in Eq. (37) to also include dependence on the depth as in Eq. (58). In the next section, we also show that this new model, when used for deconvolution, provides a much more realistic reconstruction than the conventional model.

Since the objects that we are interested in imaging in this work are essentially microspheres, their geometry is known a priori and so are the manufacture design specifications (e.g. diameter). If we denote by $\boldsymbol{\omega}_o \in \Theta$ the unknown parameter(s) of the object, and by $\boldsymbol{\omega}_h \in \Theta$ those of the PSF (especially the phase) to be estimated. Here, Θ is the parameter solution space. We will see in the next section what these individual parameters are. The knowledge of these parameters allow us to completely specify the object and the PSF.

4.1 Maximum Likelihood Approach

This was already discussed in Section 3 for the case of direct object and PSF estimation from the observation data. However, with the parametrization of the object and the PSF, the ML approach in Subsection 3.1 cannot be used in its direct form. We thus rewrite the likelihood in terms of these parameters as

$$\Pr(i|\boldsymbol{\omega}_o, \boldsymbol{\omega}_h) = \prod_{\mathbf{x}\in\Omega_s} \frac{[(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + b](\mathbf{x})^{i(\mathbf{x})} e^{-[(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + b](\mathbf{x})}}{i(\mathbf{x})!}. \tag{68}$$

By applying the -log operator to Eq. (68), we get the following energy functional

$$\mathcal{J}_{obs}(i|\boldsymbol{\omega}_o, \boldsymbol{\omega}_h) = \sum_{\mathbf{x}\in\Omega_s} [(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + b](\mathbf{x}) - \sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x}) \log[(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + b](\mathbf{x}) + \sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x})!. \tag{69}$$

Maximizing the likelihood in Eq. (68) with respect to the parameters $\boldsymbol{\omega}_o$ and $\boldsymbol{\omega}_h$ leads us to the maximum likelihood (ML) estimate of the parameters,

$$(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h) = \arg \max_{(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h)} \{\Pr(i|\boldsymbol{\omega}_o, \boldsymbol{\omega}_h)\}. \tag{70}$$

Since the term $i(\mathbf{x})!$ in Eq. (70) is independent of the parameters to be estimated, we drop it from the maximization expression. Since maximizing Eq. (68) is equivalent to minimizing

$\mathcal{J}(i|\boldsymbol{\omega}_o, \boldsymbol{\omega}_h)$, we get the following equivalent expression

$$(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h) = \arg \min_{(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h)} \mathcal{J}_{obs}(i|\boldsymbol{\omega}_o, \boldsymbol{\omega}_h). \tag{71}$$

4.2 Maximum A Posteriori Approach

As the solution space for the parameters is large, we add some realistic priors on these parameters to restrict the possible outcomes. It was shown in [51] that when the observed intensity data is available, unique reconstruction of the phase of the pupil is possible provided there is *a priori* information available on the phase or the object support. The Bayes's theorem can be used to estimate the unknown parameters $\boldsymbol{\omega}_h$ and $\boldsymbol{\omega}_o$ from the observation i. In this problem, we treat the observation, object and the PSF parameters as probability-frequency functions. Thus the conditional posterior probability is

$$\Pr(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i) = \frac{\Pr(i|\boldsymbol{\omega}_o, \boldsymbol{\omega}_h) \Pr(\boldsymbol{\omega}_o) \Pr(\boldsymbol{\omega}_h)}{\Pr(i)}, \tag{72}$$

where $\Pr(i)$ is the prior on the observation. In the above expression, we have assumed that the object and the PSF parameters are independent of each other. So each of them could be independently treated with hyperprior distributions $\Pr(\boldsymbol{\omega}_o)$ and $\Pr(\boldsymbol{\omega}_h)$. The estimates for $\boldsymbol{\omega}_o$ and $\boldsymbol{\omega}_h$ can be obtained by maximizing the joint probability or minimizing the -log of the posterior as

$$\begin{aligned} (\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h) &= \arg \max_{(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h)} \{\Pr(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i)\}, \\ &= \arg \min_{(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h)} \{-log[\Pr(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i)]\}. \end{aligned} \tag{73}$$

The parameters so obtained are known as the maximum a posteriori (MAP) estimates. The energy function in Eq. (71) can now be rewritten as the sum of three individual energy terms: the first term is obtained from the observation model in Eq. (68), the second and the third energy terms, $\mathcal{J}_{hp}(\boldsymbol{\omega}_o)$ and $\mathcal{J}_{hp}(\boldsymbol{\omega}_h)$, are generated by applying -log operator on the hyperpriors, $\Pr(\boldsymbol{\omega}_o)$ and $\Pr(\boldsymbol{\omega}_h)$.

$$(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h) = \arg \min_{(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h)} \{\mathcal{J}_{obs}(i|\boldsymbol{\omega}_o, \boldsymbol{\omega}_h) + \mathcal{J}_{hp}(\boldsymbol{\omega}_o) + \mathcal{J}_{hp}(\boldsymbol{\omega}_h)\}. \tag{74}$$

It is simple to notice that Eq. (74) reduces to Eq. (71) when the two hyperpriors are uniform. On the other hand, if we assign a Gamma prior distribution to each of the parameters, then the hyperpriors are

$$\Pr(\omega_o) = \omega_o^{\alpha_o - 1} \exp(-\beta_o \omega_o), \tag{75}$$

and

$$\Pr(\omega_h) = \omega_h^{\alpha_h - 1} \exp(-\beta_h \omega_h), \tag{76}$$

where $(\alpha_{(\cdot)}, \beta_{(\cdot)})$ are the parameters of the Gamma prior. When $(\alpha_o, \beta_o) \rightarrow 0$, then the hyperprior $\Pr(\omega_o)$ becomes the Jeffreys' noninformative prior. The generalized energy function in Eq. (74) is written by using the expressions in Eqs. (69), (75) and (76)

$$\mathcal{J}(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h | i) = \sum_{\mathbf{x}\in\Omega_s} [(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + b](\mathbf{x}) - \sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x}) \log[(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + b](\mathbf{x}) + \sum_{l=1}^{N_o} (\alpha_{o,l} - 1) \log(\omega_{o,l}) - \beta_{o,l}\omega_{o,l} + \sum_{m=1}^{N_h} (\alpha_{h,m} - 1) \log(\omega_{h,m}) - \beta_{o,m}\omega_{o,m}, \quad (77)$$

where $\alpha_{o,l}$ and $\beta_{o,l}$ are the parameters of the hyperprior for the l^{th} parameter $\omega_{o,l}$ of the object. Similarly, $\alpha_{h,m}$ and $\beta_{h,m}$ are the parameters of the hyperprior for the m^{th} parameter $\omega_{h,m}$ of the PSF. In the above equation we have assumed these parameters to be known a priori. This need not necessarily be the case.

In this MAP approach, the background fluorescence term b is determined from a single "specimen-independent" slice as $\hat{b}$. It is only estimated once for the whole volume, and is then assumed to be known during the estimation of the object $\boldsymbol{\omega}_o$ and the PSF parameters $\boldsymbol{\omega}_h$. This estimation procedure was already described in Subsection 1.1.

4.2.1 Object and Point-Spread Function Parameters Estimation

In the presence of only SA, it was mentioned in Subsection 2.2.1 that the PSF is a function of the depth d, and the refractive indices n_i and n_o. This was also shown to be the case in the PSF expressions derived in Eqs. (38) and (58). Hence, an observed source point appears shifted in depth but not in the radial plane. If the sphere is assumed to be placed at a relative position (x_o, y_o, z_o) in a given volume, then the observed image will have the centroid in the volume as (x_i, y_i, z_i), with $x_i \approx x_o$ and $y_i \approx y_o$. Another point to be noted is that due to photon loss, although uniformly distributed, the true intensity of the observed sphere s is unknown. In Subsection 1.1.3, we saw how a bandlimited object could be simulated for the experimentation and testing. In the spatial domain the object can be written as

$$o(x, y, z) = \begin{cases} 1, & \forall \sqrt{(x - x_o)^2 + (y - y_o)^2 + (z - z_o)^2} \leq R \\ 0, & \text{otherwise} \end{cases}$$

where R is the radius of the sphere specified by the manufacturer. In the above case, we have assumed the true intensity to be unity and uniformly distributed within the sphere. However, the true intensity need not necessarily be unity. Thus, the two object parameters to be estimated are the intensity and the relative depth. If we assume that the apodization function A in Eq. (29) is a constant, this term can be absorbed into the intensity of the sphere. So the combined term to be estimated will be s. In the case when the true intensity is unity, the apodization function A will be a uniform distribution of s, where $s \in \mathbb{R}$. The estimation of this entity is discussed in Subsection 4.2.2. So, for the object, the only term that needs to be estimated is the relative position z_i and hence $\boldsymbol{\omega}_o = \{z_i\}$. It was

also mentioned in Subsection 2.2.2 that for a calibrated CLSM, the variation of the PSF under different imaging conditions is essentially a factor of the following three parameters: refractive index of the objective immersion medium n_i, index of the specimen or mounting medium n_s, and depth of the specimen under the cover slip d. Thus, the PSF parameters to be estimated are $\boldsymbol{\omega}_h = \{d, n_i, n_s\}$.

Summarizing, the above discussion, the energy function in Eq. (77) can be rewritten as

$$\mathcal{J}(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h, s|i) = \sum_{\mathbf{x}\in\Omega_s} [s(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + \hat{b}](\mathbf{x}) - \sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x}) \log[s(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + \hat{b}](\mathbf{x}) + \sum_{l=1}^{N_o} (\hat{\alpha}_{o,l} - 1)\log(\omega_{o,l}) - \hat{\beta}_{o,l}\omega_{o,l} + \sum_{m=1}^{N_h} (\hat{\alpha}_{h,m} - 1)\log(\omega_{h,m}) - \hat{\beta}_{o,m}\omega_{o,m}. \quad (78)$$

The object o in the above expression is derived by normalizing the object term o in Eq. (77) so that $o : \Omega_s \mapsto [0\,1]$. As there is no closed form expression for calculation of $\boldsymbol{\omega}_o$ and $\boldsymbol{\omega}_h$ from Eq. (78), we use a Newton's optimization algorithm [52]. The Newton's method in optimization, a class of hill-climbing optimization techniques that seeks the stationary point of a function, where the gradient is 0. In this case $\mathcal{J}(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i)$ is expanded using Taylor's series expansion and approximated to the second order

$$\hat{\boldsymbol{\omega}}_o^{n+1} = \hat{\boldsymbol{\omega}}_o^n - \tau_o^n \left(\nabla^2_{\boldsymbol{\omega}_o} \mathcal{J}(\hat{\boldsymbol{\omega}}_o^n, \hat{\boldsymbol{\omega}}_h^n|i)\right)^{-1} \nabla_{\boldsymbol{\omega}_o} \mathcal{J}(\hat{\boldsymbol{\omega}}_o^n, \hat{\boldsymbol{\omega}}_h^n|i), \text{ s.t } \boldsymbol{\omega}_o \in \Theta \quad (79)$$

and

$$\hat{\boldsymbol{\omega}}_h^{n+1} = \hat{\boldsymbol{\omega}}_h^n - \tau_h^n \left(\nabla^2_{\boldsymbol{\omega}_h} \mathcal{J}(\hat{\boldsymbol{\omega}}_o^n, \hat{\boldsymbol{\omega}}_h^n|i)\right)^{-1} \nabla_{\boldsymbol{\omega}_h} \mathcal{J}(\hat{\boldsymbol{\omega}}_o^n, \hat{\boldsymbol{\omega}}_h^n|i), \text{ s.t } \boldsymbol{\omega}_h \in \Theta \quad (80)$$

where $\nabla_{\boldsymbol{\omega}_o} \mathcal{J}(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i)$ is the gradient and $\nabla^2_{\boldsymbol{\omega}_o} \mathcal{J}(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i)$ is the Hessian of the cost function with respect to the object parameter $\boldsymbol{\omega}_o$. The calculation of these two terms will be discussed in the next two Subsections. Usually Newton's method is modified to include small steps τ_o^n and τ_h^n at each iteration. This is often done to ensure that the Wolfe conditions [52] are satisfied at each step. The initial steps are chosen as $\tau_o^0 = 1$ and $\tau_h^0 = 1$ and they are updated in the following manner

$$\tau_o^{n+1} = \frac{\tau_o^n}{2}; \tau_h^{n+1} = \frac{\tau_h^n}{2} \quad (81)$$

The Eqs. (79) and (80) are alternatively carried out so that

$$\mathcal{J}(\hat{\boldsymbol{\omega}}_o^{n+1}, \hat{\boldsymbol{\omega}}_h^{n+1}|i) < \mathcal{J}(\hat{\boldsymbol{\omega}}_o^n, \hat{\boldsymbol{\omega}}_h^n|i) \quad (82)$$

The geometric interpretation of Newton's method is that at each iteration one approximates $\mathcal{J}(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h|i)$ by a quadratic function around $\hat{\boldsymbol{\omega}}_h$ or $\hat{\boldsymbol{\omega}}_o$, and then take a step towards the maximum/minimum of that quadratic function.

Gradients Calculation The general form of the energy function's partial gradient as a function of the parameters $\boldsymbol{\omega}_o$ and $\boldsymbol{\omega}_h$ can be written as

$$\begin{aligned}\nabla_{\boldsymbol{\omega}_o} J(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i) &= \sum_{\mathbf{x}\in\Omega_s} h(\mathbf{x};\boldsymbol{\omega}_h) * \frac{\partial}{\partial \boldsymbol{\omega}_o} o(\mathbf{x};\boldsymbol{\omega}_o) - \\ &\quad \frac{i(\mathbf{x})}{[(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + \hat{b}](\mathbf{x})} \cdot [h(\mathbf{x};\boldsymbol{\omega}_h) * \frac{\partial}{\partial \boldsymbol{\omega}_o} o(\mathbf{x};\boldsymbol{\omega}_o)], \forall \mathbf{x} \in \Omega_s, \quad (83)\end{aligned}$$

and

$$\begin{aligned}\nabla_{\boldsymbol{\omega}_h} J(\boldsymbol{\omega}_o, \boldsymbol{\omega}_h|i) &= \sum_{\mathbf{x}\in\Omega_s} \frac{\partial}{\partial \boldsymbol{\omega}_h} h(\mathbf{x};\boldsymbol{\omega}_h) * o(\mathbf{x};\boldsymbol{\omega}_o) - \\ &\quad \frac{i(\mathbf{x})}{[(h(\boldsymbol{\omega}_h) * o(\boldsymbol{\omega}_o)) + \hat{b}](\mathbf{x})} \cdot [\frac{\partial}{\partial \boldsymbol{\omega}_h} h(\mathbf{x};\boldsymbol{\omega}_h) * o(\mathbf{x};\boldsymbol{\omega}_o)], \forall \mathbf{x} \in \Omega_s. \quad (84)\end{aligned}$$

In Eqs. (83) and (84), we need to calculate the two derivatives $\partial o(\mathbf{x};\boldsymbol{\omega}_o)/\partial \boldsymbol{\omega}_o$ and $\partial h(\mathbf{x};\boldsymbol{\omega}_h)/\partial \boldsymbol{\omega}_h$. These two terms are derived in the Appendix C. We remind that the lateral object position does not change during the observation process, and hence the estimation of the lateral centroid of the observed bead corresponds to the true radial bead position.

Hessian Approximation Since the Hessian in Eqs. (79) and (80) are difficult to be calculated for each of the parameters, we make its approximation using the successive gradient vectors. The most efficient approach, to solve Eqs. (79) and (80) is using the Broyden-Fletcher-Goldfarb-Shanno or BFGS method (see Chapter 8 of [52]). We will use this method here as well, leaving the implementation details to [52].

4.2.2 Uniform Intensity and Apodization Estimation

We discuss here the estimation of the factor s, which is a combination of the intensity and the apodization function. We have assumed that the apodization function is a constant. Since the energy function in Eq. (78) is convex in s, minimizing the energy function with respect to s is equivalent to equating the gradient $\nabla_s \mathcal{J}(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h, s|i)$ to zero. That is

$$\begin{aligned}\nabla_s \mathcal{J}(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h, s|i) &= \frac{\partial}{\partial s}\{\sum_{\mathbf{x}\in\Omega_s} [s(h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)) + \hat{b}](\mathbf{x}) - \\ &\quad \sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x}) \log[s(h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)) + \hat{b}](\mathbf{x})\} = 0 \quad (85)\end{aligned}$$

where the terms independent of s were dropped from Eq. (78), and the parameters $\boldsymbol{\omega}_o$ and $\boldsymbol{\omega}_o$ are assumed to be known or prior estimated.

$$\begin{aligned}\nabla_s \mathcal{J}(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h, s|i) &= \sum_{\mathbf{x}\in\Omega_s} [h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)](\mathbf{x}) - \\ &\quad \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x})}{[s(h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)) + \hat{b}](\mathbf{x})} \cdot [h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)](\mathbf{x}) \\ &= 0 \end{aligned} \tag{86}$$

and hence

$$\sum_{\mathbf{x}\in\Omega_s} [h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)](\mathbf{x}) = \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x})}{[s(h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)) + \hat{b}](\mathbf{x})} \cdot [h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)](\mathbf{x}), \forall \mathbf{x} \in \Omega_s. \tag{87}$$

This would be possible only if

$$1(\mathbf{x}) = \frac{i(\mathbf{x})}{[s(h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)) + \hat{b}](\mathbf{x})}, \forall \mathbf{x} \in \Omega_s \tag{88}$$

or equivalently

$$s = \frac{1}{(N_x N_y N_z)} \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x}) - \hat{b}(\mathbf{x})}{[h(\hat{\boldsymbol{\omega}}_h) * o(\hat{\boldsymbol{\omega}}_o)](\mathbf{x})}, \forall \mathbf{x} \in \Omega_s \tag{89}$$

In an iterative form, the above expression for calculating s can now be written as

$$\hat{s}^{n+1} = \frac{1}{(N_x N_y N_z)} \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x}) - \hat{b}(\mathbf{x})}{[h(\hat{\boldsymbol{\omega}}_h^n) * o(\hat{\boldsymbol{\omega}}_o^n)](\mathbf{x})}. \tag{90}$$

5 Implementation and Analysis

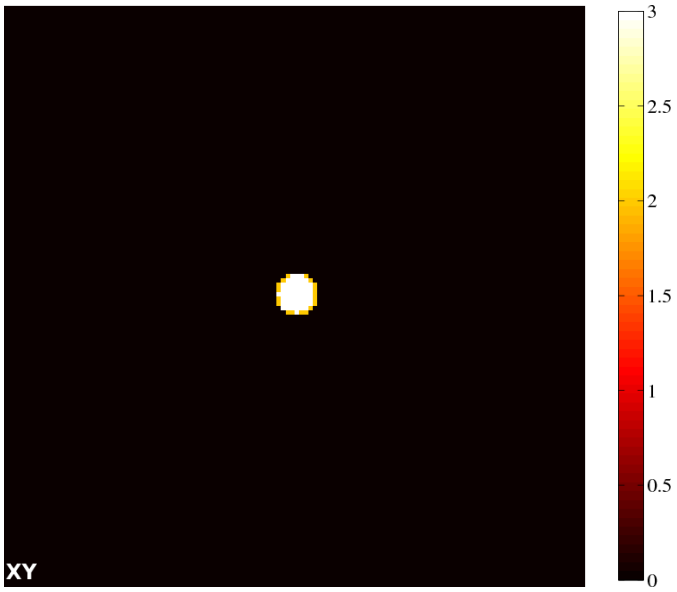
For the implementation, a bandlimited object was simulated as explained in Subsection 1.1.3. The radius of the microsphere was chosen to be 250nm, and it is assumed to be embedded in a medium with $n_s = 1.33$. The objective lens used is a '*Plan-Neofluar*' oil immersion lens ($n_i = 1.518$) with 40X magnification and Numerical Aperture (NA) of 1.3. The excitation and emission peaks are at wavelengths of 488nm (λ_{ex}) and 520nm (λ_{em}) respectively. The physical pinhole size was fixed at 61μm, and the images were sampled at lateral and axial pixel sizes of 46.92nm and 166.16nm. Fig. 13 shows the simulated bandlimited microsphere, the PSF calculated from Eq. (60) at a depth of 5μm, and the simulated observation (with $\gamma = 100$). We denote by $\boldsymbol{\omega}_{h,true} = \{1.518, 1.33, 10\mu m\}$, the true experimental settings of the simulation.

5.1 Initialization of the Algorithm

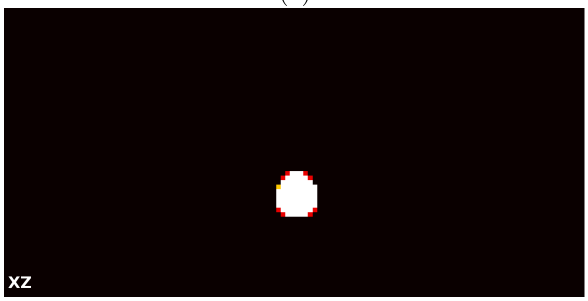
The microsphere is often not laterally and axially centered so its approximate relative position (x_0, y_0, z_0) has to be calculated from the observed images. We propose a simple approach in estimating the relative microsphere position in the observed images. The only assumption made here is that the observation data has been treated so that the available data is from a single microsphere. By knowing the physical diameter of the microsphere (here 500nm), we can locate its centroid in the observed volume by estimating the center of gravity (COG) of the intensities [53]. The above method gives very precise lateral locations ($\hat{x}_0$, $\hat{y}_0$). For example, Fig. 16 shows segmentation of the observation data and a precise radial COG estimation was obtained from this segmentation. However, due to focal anomaly the axial relative location ($\hat{z}_0$) of the object cannot be accurately obtained from the observation data. In the example given above, the COG of the observation (the actual focal position (AFP)) in the axial case was accurately estimated as 8 slices off the central plane. This does not at all correspond with the true position of the object (nominal focal position (NFP)) in the volume. From geometrical optics [54], if the objective is nonparaxial, the estimated axial position from the observation ($\hat{z}_0$) is multiplied by a factor $(\tan(\arcsin(NA/n_i^{(0)}))/(\tan(\arcsin(NA/n_s^{(0)})))$ to get the new relative position ($\hat{z}_{0,\text{new}}$). While for the paraxial case, the multiplication factor becomes $(n_i^{(0)}/n_s^{(0)})$. These new values are then used as starting points for the iteration. It was found that the initialization process had a maximum error of 1% in estimating the lateral positions and about 8.4% in estimating the axial position. The parameters of the PSF $\hat{\boldsymbol{\omega}}_h^{(0)}$ are assigned valid non zero values as starting estimates.

5.2 Preliminary Results

In this section we discuss the preliminary results that was obtained by the parameter estimation algorithm. Fig. 17 shows the plot of the change in the cost function and the different parameters with iteration. The refractive indices n_i and n_s were estimated to be 1.5163



(a)



(b)

Figure 16: Segmentation of the MIP along the (a) optic axis projection giving radial segment and (b) lateral projection giving axial segment (© Ariana-INRIA/I3S).

and 1.3201 which are very close to their true values of 1.518 and 1.33 chosen for simulation. Although two parameters were accurately determined with a maximum error of $< 1\%$, there were errors in estimating the relative true position of the object in the volume and the nominal focal depth d. The relative position of the object in the volume was found to be about 4 slices below the central plane, and differs from the true position by about 2 slices. Due to this reason, the cost function does not reach its global minimum as seen in Fig. 17(a).
In Fig. 18, we have shown the ℓ_∞ normalized axial intensity profiles for the true object, observation, restoration with spherically aberrated approximation and restoration with

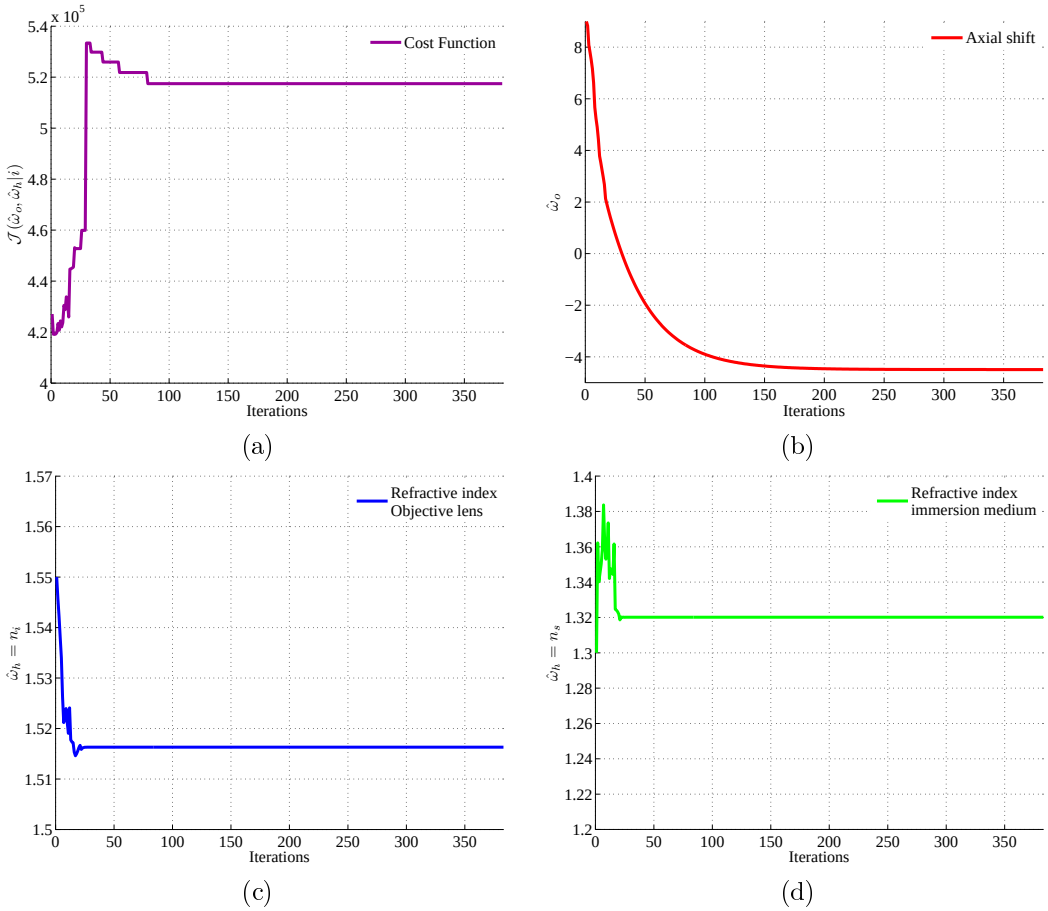


Figure 17: Graph showing the progression of the (a) cost function $\mathcal{J}(\hat{\boldsymbol{\omega}}_o, \hat{\boldsymbol{\omega}}_h|i)$, (b) relative position $\hat{\boldsymbol{\omega}}_o$, (c) objective lens refractive index $\hat{n}_i$, and (d) specimen medium refractive index $\hat{n}_s$ estimation with iterations (© Ariana-INRIA/I3S).

diffraction-limited approximation. It can be seen that the observed intensity profile is axially shifted from the true axial position by about 8 slices. After restoration with the spherically aberrated PSF, the estimated object's axial location corresponds very well with the true position of the object. The intensity profile is quite symmetrical in comparison to the observed profile. However, when restoration is done using a diffraction-limited approximation (ignoring aberrations) has lead to error in locating the actual axial position. We can also observe that the intensity is highly asymmetrical and is worse than the observed profile although the restoration is quite free of blurring.

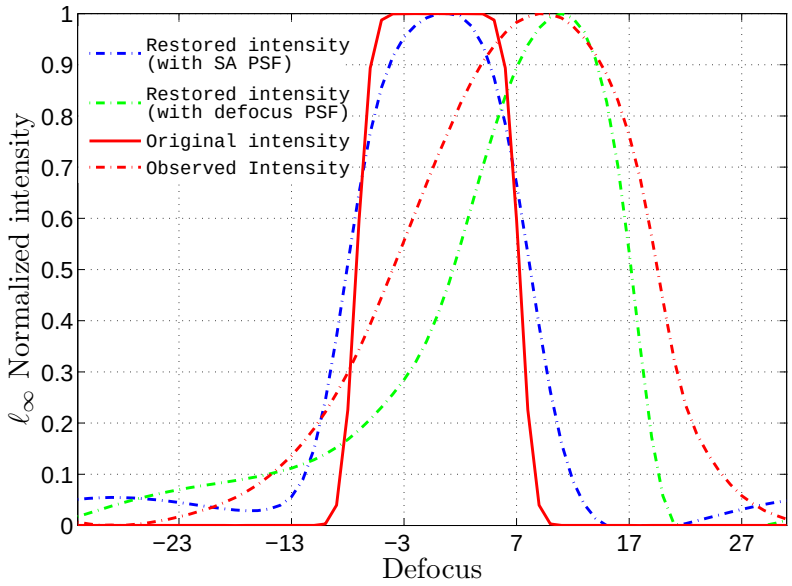


Figure 18: Axial intensity profiles for the true object, the observed object, the restored intensity with a diffraction-limited PSF, and the restored intensity with a spherically aberrated PSF. The intensities are ℓ_∞ normalized for visual comparison (© Ariana-INRIA/I3S).

6 Conclusions and Future Work

In this research report, we proposed an approach for estimating the PSF from an observation data given some knowledge of the object. As a first step we have restricted the observed object to a specific geometry. The validation of the algorithm on simulated data shows very promising results for the problem of PSF extraction from observed intensities for a fluorescence microscope where spherical aberration is the dominant form of aberration. In the case, where a sub-resolution microsphere is used, bead size correction will not be necessary and the object should be treated as a Dirac function. Future work is aimed at testing the proposed approach on images of fluorescent polystyrene latex Tetraspeck™microspheres from Invitrogen™. One possible extension of this work might involve applying this method to restoring biological specimens affected by spherical aberrations adding some constraints on the object (spatial or frequency). The task is not simple as there are many possible solutions for the phase function, though a realization might be possible through regularization. This work also opens up new possibilities into the field of depth-varying image restoration.

A Appendix: Maximum Likelihood Expectation Maximization (MLEM)

In this appendix, we summarize the MLEM algorithm as derived by Richardson [23]. Given the observation $i(\mathbf{x})$ and the PSF $h(\mathbf{x})$, when the requirement is only to find the original object $o(\mathbf{x})$, we use the Bayes theorem

$$\Pr(o(\mathbf{x})|i(\mathbf{x}')) = \frac{\Pr(i(\mathbf{x}')|o(\mathbf{x}))\Pr(o(\mathbf{x}))}{\sum_{\mathbf{x}''\in\Omega_s}\Pr(i(\mathbf{x})|o(\mathbf{x}''))\Pr(o(\mathbf{x}''))}; \mathbf{x},\mathbf{x}',\mathbf{x}''\in\Omega_s. \tag{91}$$

Considering all the $i(\mathbf{x}')$ and its dependance on $o(\mathbf{x})$, we get

$$\Pr(o(\mathbf{x})) = \sum_{\mathbf{x}'\in\Omega_s}\Pr(o(\mathbf{x})i(\mathbf{x}')) = \sum_{\mathbf{x}'\in\Omega_s}\Pr(o(\mathbf{x})|i(\mathbf{x}'))\Pr(i(\mathbf{x}')), \tag{92}$$

since $Pr(o(\mathbf{x})|i(\mathbf{x}')) = \Pr(o(\mathbf{x})i(\mathbf{x}'))/\Pr(i(\mathbf{x}'))$. Substituting Eq. (91) in Eq. (92) we can say

$$\Pr(o(\mathbf{x})) = \sum_{\mathbf{x}'\in\Omega_s}\frac{\Pr(i(\mathbf{x}')|o(\mathbf{x}))\Pr(o(\mathbf{x}))\Pr(i(\mathbf{x}'))}{\sum_{\mathbf{x}''\in\Omega_s}\Pr(i(\mathbf{x})|o(\mathbf{x}''))\Pr(o(\mathbf{x}''))} \tag{93}$$

Since the left and the right hand side of Eq. (93) have the desired result $\Pr(o(\mathbf{x}))$, we write the result in the following iterative form

$$\Pr{}^{(n+1)}(o(\mathbf{x})) = \Pr{}^{(n)}(o(\mathbf{x}))\sum_{\mathbf{x}'\in\Omega_s}\frac{\Pr(i(\mathbf{x}')|o(\mathbf{x}))\Pr(i(\mathbf{x}'))}{\sum_{\mathbf{x}''\in\Omega_s}\Pr(i(\mathbf{x}')|o(\mathbf{x}''))\Pr^{(n)}(o(\mathbf{x}''))}, n=\{0,1,\ldots\} \tag{94}$$

where the initial value $\Pr^{(0)}(o(\mathbf{x}))$ is assumed to be estimated. Using Bayes postulate a uniform distribution is assumed for the initial estimate so that $\Pr^{(0)}(o(\mathbf{x})) = o^{(0)}(\mathbf{x})/\sum_{\mathbf{x}\in\Omega_s}o(x)$. Also, $\Pr(o(\mathbf{x})) = o(\mathbf{x})/\sum_{\mathbf{x}\in\Omega_s}o(x)$ and $\Pr(i(\mathbf{x})) = i(\mathbf{x})/\sum_{\mathbf{x}\in\Omega_s}i(x) = i(\mathbf{x})/\sum_{\mathbf{x}\in\Omega_s}o(x)$, by the property of flux conservation during the restoration and $\sum_{\mathbf{x}\in\Omega_s}i(x) = \sum_{\mathbf{x}\in\Omega_s}o(x)$. Similarly, $\Pr(i(\mathbf{x}')|o(\mathbf{x})) = \Pr(h(\mathbf{x}'-\mathbf{x})) = h((\mathbf{x}'-\mathbf{x}))/\sum_{\mathbf{x}''\in\Omega_s}h(\mathbf{x}'')$. So, Eq. (94) becomes

$$\left(\frac{o^{(n+1)}(\mathbf{x})}{\sum_{\mathbf{x}\in\Omega_s}o(\mathbf{x})}\right) = \left(\frac{o^{(n)}(\mathbf{x})}{\sum_{\mathbf{x}\in\Omega_s}o(\mathbf{x})}\right)\sum_{\mathbf{x}'\in\Omega_s}\frac{\left(h(\mathbf{x}'-\mathbf{x})/\sum_{\mathbf{x}''\in\Omega_s}h(\mathbf{x}'')\right)\cdot\left(i(\mathbf{x}')/\sum_{\mathbf{x}\in\Omega_s}o(\mathbf{x})\right)}{\sum_{\mathbf{x}''\in\Omega_s}\left(h(\mathbf{x}'-\mathbf{x}'')/\sum_{\mathbf{x}''\in\Omega_s}h(\mathbf{x}'')\right)\cdot\left(o^{(n)}(\mathbf{x}'')/\sum_{\mathbf{x}\in\Omega_s}o(\mathbf{x})\right)} \tag{95}$$

or

$$o^{(n+1)}(\mathbf{x}) = o^{(n)}(\mathbf{x})\sum_{\mathbf{x}'\in\Omega_s}\frac{h(\mathbf{x}'-\mathbf{x})\cdot i(\mathbf{x}')}{\sum_{\mathbf{x}''\in\Omega_s}h(\mathbf{x}'-\mathbf{x}'')\cdot o^{(n)}(\mathbf{x}'')}, n=\{0,1,\ldots\} \tag{96}$$

B Appendix: Fourier transform of a Gaussian function

A function of three independent variables is called *separable* with respect to a specific co-ordinate system if it can be written as a product of three functions, each of which depends only on one of the independent variables. Thus the function f is separable in coordinates (x, y, z) if

$$f(x, y, z) = f_X(x) f_Y(y) f_Z(z). \tag{97}$$

The 3D Gaussian function $h : \mathbb{N}^3 \rightarrow \mathbb{R}_+$ is considered to be separable as the function can be written individually as the combination of several other functions. The 3D convolution of h with $o : \mathbb{N}^3 \rightarrow \mathbb{R}_+$ is thus reduced to three successive 1D multiplications in the Fourier domain with the Fourier transform of o. Thus,

$$(h * o)(x, y, z) = (h(x) * (h(y) * (h(z) * o(x, y, z)))), \tag{98}$$

and

$$(h * o)(x, y, z) = \mathcal{F}^{-1}(\mathcal{F}(h(z)) \cdot (\mathcal{F}(h(y)) \cdot (\mathcal{F}(h(x)) \cdot \mathcal{F}(o(x, y, z))))). \tag{99}$$

The Fourier transform of the continuous function h is given by:

$$\mathcal{F}(h(x)) \equiv H(k_x) = \int_{-\infty}^{+\infty} h(x) e^{-j2\pi k_x x} dx, \tag{100}$$

where $j^2 = -1$, $h(x) = (1/\sqrt{2\pi}\sigma_x) \exp(-x^2/(2\sigma_x^2))$ is the 1D Gaussian function, and k_x is the co-ordinate in the frequency domain. A closed form expression for $H(k_x)$ exists and the analytical expression can be written as

$$\begin{aligned} H(k_x) &= \int_{-\infty}^{+\infty} \frac{1}{(\sqrt{2\pi}\sigma_x)} \exp(-\frac{x^2}{(2\sigma_x^2)}) \exp(-j2\pi k_x x) dx \\ &= \exp\Big(-\frac{(2\pi k_x \sigma_x)^2}{2}\Big). \end{aligned} \tag{101}$$

From the above expression, it is clear that the Fourier transform of a Gaussian function is also Gaussian. The proof is not straight forward and is given below.

Proof. It is simple to show that differentiation of a 1D Gaussian function $h(x)$ with respect to x is

$$\frac{d}{dx} h(x) = \left(\frac{-x}{\sigma_x^2}\right) h(x). \tag{102}$$

Applying the Fourier transform to the above equation, and by using the differentiation property of the Fourier Transform, we get

$$\begin{aligned}(j2\pi k_x)H(k_x) &= \frac{-1}{\sigma_x^2}\int_{-\infty}^{+\infty} xh(x)\exp(-j2\pi k_x x)dx, \\ &= \frac{-j}{\sigma_x^2}\frac{d}{dk_x}H(k_x). \end{aligned} \tag{103}$$

Eq. (103) can be simplified to

$$\frac{1}{H(k_x)}\frac{d}{dk_x}H(k_x) = -2\pi k_x\sigma_x^2. \tag{104}$$

Integrating both sides of Eq. (104) with respect to k_x, we get

$$\int_0^{k_x}\frac{1}{H(k'_x)}\frac{d}{dk'_x}H(k'_x)dk'_x = -2\pi\sigma_x^2\int_0^{k_x}k'_x dk'_x. \tag{105}$$

Or equivalently

$$\log(H(k_x)) - \log(H(0)) = -\frac{\sigma_x^2}{2}(2\pi\sigma_x)^2. \tag{106}$$

Thus, we arrive at the expression:

$$H(k_x) = H(0)\exp(-\frac{\sigma_x^2(2\pi k_x)^2}{2}). \tag{107}$$

Since $H(0) = 1$, we arrive at Eq. (101). □

Another method to derive the above transform but using the Laplace integral is given by K. Kim and G. Shevlyakov in [55].

C Appendix: Gradient Calculations

C.1 For the Object Function

If the bead is not axially centered in the observed volume, then the relative true position $z_0 \neq 0$ and

$$o(\mathbf{x};\boldsymbol{\omega}_o) \equiv o(x, y, z - z_o), \tag{108}$$

It can be shown that in the Fourier domain this translates to

$$o(x, y, z - z_o) = o(x, y, z) * \delta(0, 0, z_o) = \mathcal{F}^{-1}\{O(\mathbf{k})e^{-j\frac{2\pi}{N_z}k_z z_o}\} \tag{109}$$

So, $\partial o(\mathbf{x};\boldsymbol{\omega}_o)/\partial\boldsymbol{\omega}_o$ becomes

$$\frac{\partial}{\partial\boldsymbol{\omega}_o}o(\mathbf{x};\boldsymbol{\omega}_o) = \frac{\partial}{\partial z_o}o(x, y, z - z_o) = -\frac{j2\pi}{N_z}\mathcal{F}^{-1}\{k_z O(\mathbf{k})e^{-j\frac{2\pi}{N_z}k_z z_o}\}. \tag{110}$$

By using Eqs. (83) and (110), we can calculate the gradients with respect to $\boldsymbol{\omega}_o$ required at every iteration n of Eq. (79). However, the calculations for the PSF parameters in Eq. (80) are not so straightforward.

C.2 For the Point-Spread Function

The gradient of the theoretically calculated PSF in Eq. (60) can be written as

$$\begin{aligned}
\frac{\partial}{\partial\boldsymbol{\omega}_h}h_{Th}(\mathbf{x};\lambda_{ex},\lambda_{em}) &= \frac{\partial}{\partial\boldsymbol{\omega}_h}\left\{|h_A(\mathbf{x};\lambda_{ex})|^2 \cdot \int\limits_{x_1^2+y_1^2\leq\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{em})|^2 dx_1 dy_1\right\} \\
&= |h_A(\mathbf{x};\lambda_{ex})|^2 \cdot \frac{\partial}{\partial\boldsymbol{\omega}_h}\left\{\int\limits_{x_1^2+y_1^2\leq\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{em})|^2 dx_1 dy_1\right\} + \\
&\quad \frac{\partial}{\partial\boldsymbol{\omega}_h}\left\{|h_A(\mathbf{x};\lambda_{ex})|^2\right\} \cdot \int\limits_{x_1^2+y_1^2\leq\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{em})|^2 dx_1 dy_1 \\
&= |h_A(\mathbf{x};\lambda_{ex})|^2 \cdot \int\limits_{x_1^2+y_1^2\leq\frac{D^2}{4}} \frac{\partial}{\partial\boldsymbol{\omega}_h}\left\{|h_A(\mathbf{x};\lambda_{em})|^2\right\} dx_1 dy_1 + \\
&\quad \frac{\partial}{\partial\boldsymbol{\omega}_h}\left\{|h_A(\mathbf{x};\lambda_{ex})|^2\right\} \cdot \int\limits_{x_1^2+y_1^2\leq\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{em})|^2 dx_1 dy_1
\end{aligned} \tag{111}$$

For Eq. (111), it is sufficient to calculate $\partial h_A(\mathbf{x};\lambda)/\partial \boldsymbol{\omega}_h$ for a general wavelength λ and then use the same functions for λ_{ex} and λ_{em}. Thus

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\omega}_h}|h_A(\mathbf{x};\lambda)|^2 &= \frac{\partial}{\partial \boldsymbol{\omega}_h}\{h_A(\mathbf{x};\lambda)h_A^*(\mathbf{x};\lambda)\} \\
&= \frac{\partial}{\partial \boldsymbol{\omega}_h}\{h_A(\mathbf{x};\lambda)\}h_A^*(\mathbf{x};\lambda) + h_A(\mathbf{x};\lambda)\frac{\partial}{\partial \boldsymbol{\omega}_h}\{h_A^*(\mathbf{x};\lambda)\} \\
&= \left(h_A^*(\mathbf{x};\lambda)\frac{\partial}{\partial \boldsymbol{\omega}_h}\{h_A(\mathbf{x};\lambda)\}\right) + \left(h_A^*(\mathbf{x};\lambda)\frac{\partial}{\partial \boldsymbol{\omega}_h}\{h_A(\mathbf{x};\lambda)\}\right)^* \quad (112)
\end{aligned}$$

Assuming the apodization function to be unity, the derivative of the amplitude PSF can be written as

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\omega}_h}h_A(\mathbf{x};\lambda) &= \frac{\partial}{\partial \boldsymbol{\omega}_h}\mathcal{F}^{-1}\{\exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\} \\
&= \mathcal{F}^{-1}\{\frac{\partial}{\partial \boldsymbol{\omega}_h}\exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\} \\
&= \mathcal{F}^{-1}\{jk_0\exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\frac{\partial}{\partial \boldsymbol{\omega}_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\} \\
&= jk_0\mathcal{F}^{-1}\{\exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\frac{\partial}{\partial \boldsymbol{\omega}_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\} \quad (113)
\end{aligned}$$

But,

$$\begin{aligned}
\varphi(\theta_i,\theta_s,z;d,n_i,n_s) &= \varphi_d(\theta_s,z;d,n_i,n_s) + \varphi_a(\theta_i,\theta_s,z;d,n_i,n_s), \quad (114) \\
&= n_s z(1-\cos\theta_s) + d(n_s\cos\theta_s - n_i\cos\theta_i). \quad (115)
\end{aligned}$$

It is straight forward to show that

$$\left[\frac{\partial}{\partial \boldsymbol{\omega}_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s)\right]_{\boldsymbol{\omega}_h=d} = (n_s\cos\theta_s - n_i\cos\theta_i), \quad (116)$$

and also

$$\begin{aligned}
\left[\frac{\partial}{\partial \boldsymbol{\omega}_h}n_s\cos\theta_s\right]_{\boldsymbol{\omega}_h=n_s} &= \frac{\partial}{\partial \boldsymbol{\omega}_h}n_s\frac{\sqrt{k_s^2-(k_x^2+k_y^2)}}{k_s} \\
&= \frac{\partial}{\partial \boldsymbol{\omega}_h}\sqrt{n_s^2-n_k^2}, \text{ where } n_k = \frac{\lambda}{2\pi}\sqrt{k_x^2+k_y^2} \\
&= \frac{n_s}{\sqrt{n_s^2-n_k^2}} \\
&= \frac{k_s}{\sqrt{k_s^2-(k_x^2+k_y^2)}} \\
&= \sec\theta_s \quad (117)
\end{aligned}$$

From the Eq. (117), we can say

$$\left[\frac{\partial}{\partial \boldsymbol{\omega}_h}\varphi(\theta_i, \theta_s, z; d, n_i, n_s)\right]_{\boldsymbol{\omega}_h = n_i} = -d\sec\theta_i, \tag{118}$$

and

$$\left[\frac{\partial}{\partial \boldsymbol{\omega}_h}\varphi(\theta_i, \theta_s, z; d, n_i, n_s)\right]_{\boldsymbol{\omega}_h = n_s} = z(1-\sec\theta_s) + d\sec\theta_s. \tag{119}$$

Summarizing, the gradient of a theoretical PSF model $h_{\mathrm{Th}}(\mathbf{x}; \boldsymbol{\omega}_h)$ with respect to its parameters $\boldsymbol{\omega}_h$, can be obtained by combining the Eqs. (111), (112), (113), (116), (118) and (119).

D Appendix: Emperical PSF Imaging

There are several methods available in literature for experimentally imaging microspheres [19, 43]. However, for our experiments, we are interested in artificially simulating the mismatch in refractive indices between the objective lens and the specimen medium. In order to achieve this, we have used a simple technique where the microspheres are stuck to the bottom of a cover slip. The coverslip is placed in water and imaged using a water immersion lens. In this way the depth is fixed and there are no additional aberrations. Fig. 19 shows the schematic of the experiment. In this case, the aberrations are caused due to mismatch in index between n_i and n_g. The use of very small beads lead to a bad SNR, so that one

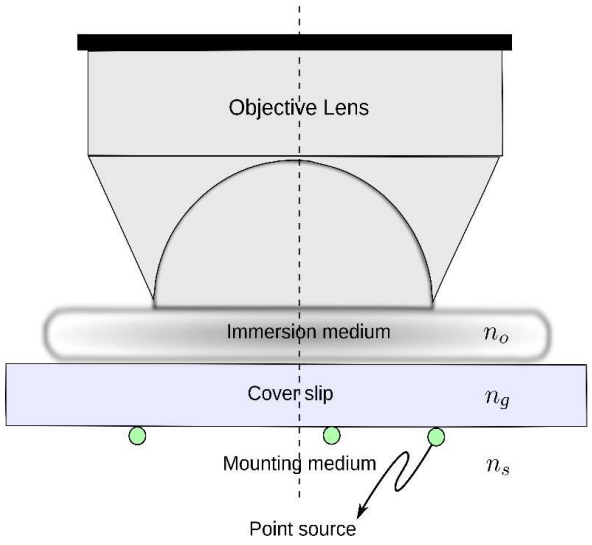


Figure 19: Schematic describing the experimental procedure for generating spherical aberration due to refractive index mismatch when imaging point sources (© Ariana-INRIA/I3S).

often uses beads of the order of the microscope resolution. We have used polystyrene latex microspheres from Tetraspeck™with a diameter of about 170nm and having a peak excitation/emission wavelengths of 505/515 nm (green). The imaged microspheres are over sized but it does not pose a problem as we have a method for compensating and estimating finally the PSF. The coverslip has a refractive index of about 1.522 and thickness 170μ m. The objective used is a '*C-Apochromat*' of 63X magnification and NA 1.2. The correction collar is placed in the neutral position to avoid any aberration compensation and the pinhole size is 112μ m. Fig. 20 shows the microspheres distributed at various positions in the radial plane. Since they are all stuck to the bottom of the cover slip, they are at the same depth. The images are sampled at a radial sampling for 0.037μ m and axial sampling of 0.151μ m. From the ZY section in Fig. 20, we see that two beads that are positioned at the same depth

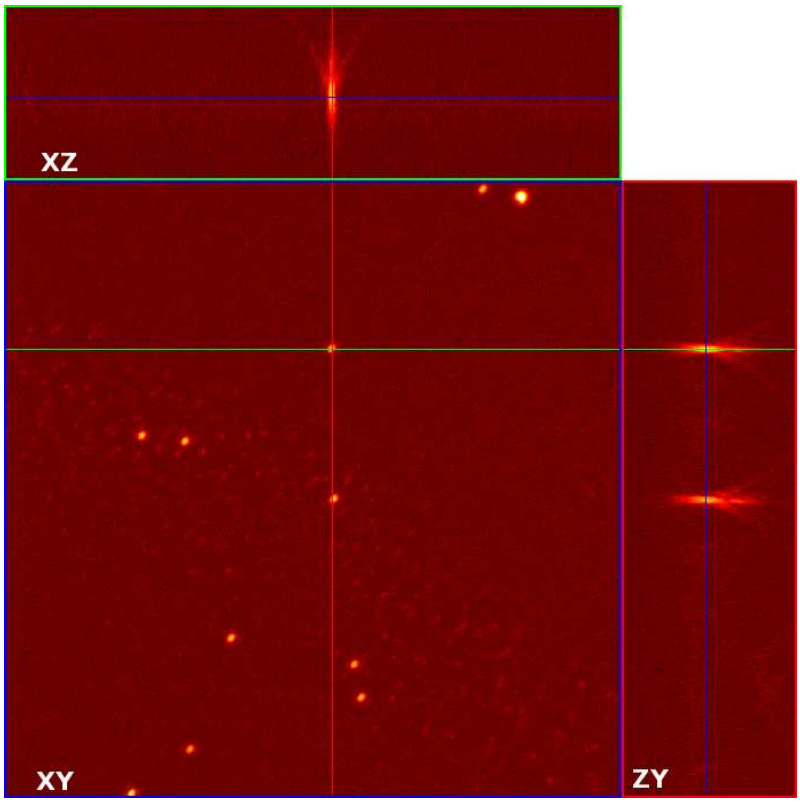
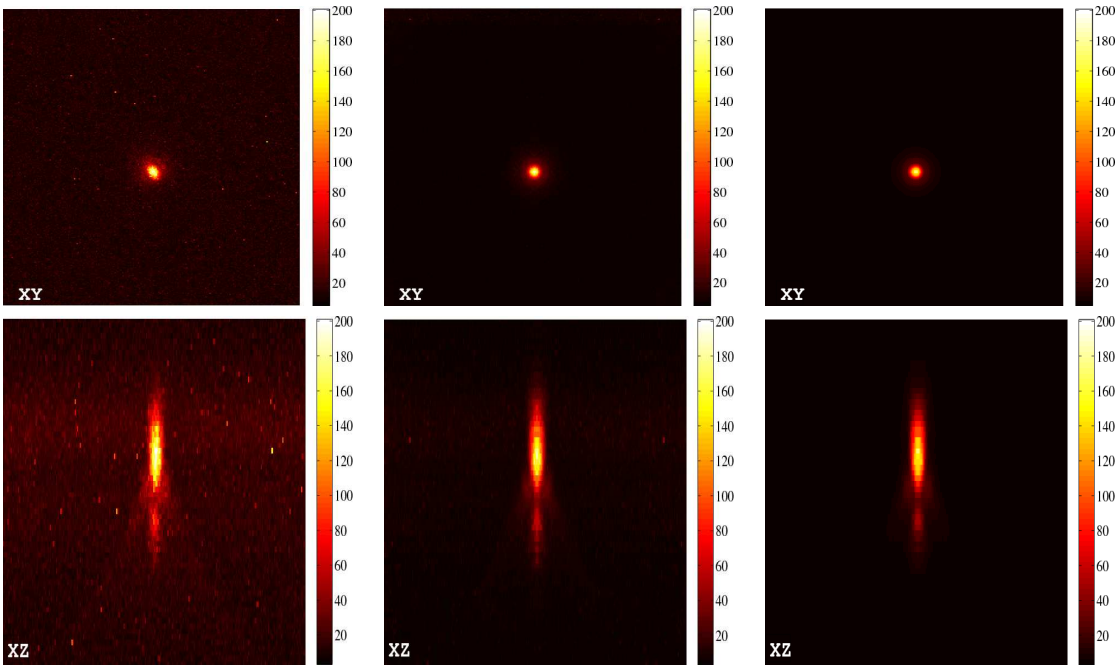


Figure 20: Observed images of microspheres that are stuck to the bottom of a cover slip demonstrates radial invariance property (© INRA Sophia-Antipolis).

but at different radial position are very similar. This validates our assumption that the PSF is approximately invariant to radial translations.

Fig. 21(a) shows the radial and axial maximum intensity projections of the cropped observation of a single 170nm microsphere stuck on to a 1.5 type cover slip. Since the objective used was a water immersion lens, and we are imaging into a cover slip, the resulting observed PSF was axially asymmetrical. This aberrated bead volumes were then radially averaged and Fig. 21(b) shows the corresponding MIP of the radial and axial planes. Fig. 21(c) show the stacks of 2D planes that were individually denoised using a TV scheme [56].



(a) (b) (c)

Figure 21: Maximum intensity projection (MIP) along the (a) optic axis gives the lateral plane (top) and radial axis gives axial plane (bottom) of an observed microbead; (b) the planes after circular averaging the data (a); (c) after denoising the circularly averaged data in (b). (© Ariana-INRIA/I3S, and INRA Sophia-Antipolis).

References

[1] M. Minsky. Memoir on inventing the confocal scanning microscope. *Scanning*, 10:128–138, 1988.

[2] J. B. Pawley, editor. *Handbook of Biological Confocal Microscopy*. Springer, 3^{rd} edition, 2006.

[3] L. Mandel. Sub-Poissonian photon statistics in resonance fluorescence. *Opt. Lett.*, 4:205–207, July 1979.

[4] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley, Reading, Mass., 1987.

[5] G. M. P. Van Kempen and L. J. Van Vliet. Background estimation in nonlinear image restoration. *J. Opt. Soc. Am. A*, 17(3):425–433, 2000.

[6] B. M. Hanser, M. G. Gustafsson, D. A. Agard, and J. W. Sedat. Phase retrieval for high-numerical-aperture optical systems. *Opt. Lett.*, 28(10):801–803, May 2003.

[7] P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. Parametric blind deconvolution for confocal laser scanning microscopy. In *Proc. IEEE Eng. Med. Biol. Soc.*, pages 6531–6534, Lyon, France, August 2007.

[8] P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. Parametric blind deconvolution for confocal laser scanning microscopy (CLSM)-proof of concept. Research Report 6493, INRIA, Sophia-Antipolis, France, March 2008.

[9] P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. Blind deconvolution for diffraction-limited fluorescence microscopy. In *Proceedings of IEEE International Symposium on Biomedical Imaging*, pages 740–743, Paris, France, May 2008.

[10] P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. Blind deconvolution for thin-layered confocal imaging. *Appl. Opt.*, 48(22):4437–4448, 2009.

[11] B. Zhang, J. Zerubia, and J. C. Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt.*, 46(10):1819–1829, 2007.

[12] A. Egner, M. Schrader, and S. W. Hell. Refractive index mismatch induced intensity and phase variations in fluorescence confocal, multiphoton and 4Pi-microscopy. *Opt. Comm.*, 153:211–217, August 1998.

[13] X. Lai, Z. Lin, E. S. Ward, and R. J. Ober. Noise suppression of point spread functions and its influence on deconvolution of three-dimensional fluorescence microscopy image sets. *J. Microsc.*, 217(1):93–108, 2005.

[14] P. A. Stokseth. Properties of a defocused optical system. *J. Opt. Soc. Am. A*, 59:1314–1321, October 1969.

[15] M. Born and E. Wolf. *Principles of Optics*. Cambridge U. Press, 1999.

[16] J. W. Goodman. *Introduction to Fourier Optics*. Roberts & Company Publishers, 2004.

[17] H. H. Hopkins. The frequency response of a defocused optical system. *Proc. R. Soc. London Ser. A.*, 231:91–103, July 1955.

[18] J. W. Shaevitz and D. A. Fletcher. Enhanced three-dimensional deconvolution microscopy using a measured depth-varying point-spread function. *J. Opt. Soc. Am. A*, 24(9):2622–2627, 2007.

[19] S. F. Gibson and F. Lanni. Diffraction by a circular aperture as a model for three-dimensional optical microscopy. *J. Opt. Soc. Am. A*, A6:1357–1367, 1989.

[20] M. J. Booth, M. A. A. Neil, and T. Wilson. Aberration correction for confocal imaging in refractive-index-mismatched media. *J. Microsc.*, 192:90–98, 1998.

[21] A. S. Carasso. Direct blind deconvolution. *SIAM J. Appl. Math*, 61(6):1980–2007, 2001.

[22] P. Sarder and A. Nehorai. Deconvolution methods for 3-D fluorescence microscopy images. *IEEE Signal Process. Mag.*, 23(3):32–45, May 2006.

[23] W. H. Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am. A*, 62(1):55–59, January 1972.

[24] T. J. Holmes. Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach. *J. Opt. Soc. Am. A*, 9(7):1052–1061, 1992.

[25] R. W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of the phase from image and diffraction plane pictures. *Optik*, 35:237–246, 1972.

[26] D. S. C. Biggs and M. Andrews. Iterative blind deconvolution of extended objects. In *Proc. IEEE Int. Conf. Image Process.*, volume 2, pages 454–457, October 1997.

[27] Y. Weiss and W. T. Freeman. What makes a good model of natural images? In *Proc. IEEE Comp. Vis. and Pat. Recog.*, pages 1–8, Minneapolis, Minnesota, USA, June 2007.

[28] N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, P. Roux, J.-C Olivo-Marin, and J. Zerubia. Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microsc. Res. Tech.*, 69:260–266, 2006.

[29] E. F. Y. Hom, F. Marchis, T. K. Lee, S. Haase, D. A. Agard, and J. W. Sedat. AIDA: an adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data. *J. Opt. Soc. Am. A*, 24(6):1580–1600, 2007.

[30] P. J. Verveer, M. J. Gemkow, and T. M. Jovin. A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy. *J. Microsc.*, 193:50–61, 1999.

[31] M.A.T. Figueiredo and R. D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.*, 12(8):906–916, August 2003.

[32] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *Proc. ACM Trans. on Graphics*, 25:787–794, 2006.

[33] S. Alenius and U. Ruotsalainen. Bayesian image reconstruction for emission tomography based on median root prior. *Eur. J. Nucl. Med. Mol. Imaging*, 24(3):258–265, March 1997.

[34] T. Kenig, Z. Kam, and A. Feuer. Blind image deconvolution using machine learning for three-dimensional microscopy. *IEEE Trans. PAMI*, 2009. submitted for review.

[35] J. Markham and J.-À. Conchello. Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur. *J. Opt. Soc. Am. A*, 16(10):2377–2391, 1999.

[36] P. Török, P. Varga, Z. Laczik, and G. R. Booker. Electromagnetic diffraction of light focused through a planar interface between materials of mismatched refractive indices: an integral representation. *J. Opt. Soc. Am. A*, 12(2):325–332, 1995.

[37] D. Kundur and D. Hatzinakos. Blind image restoration via recursive filtering using deterministic constraints. In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, volume 4, pages 2283–2286, May 1996.

[38] D. Kundur and D. Hatzinakos. A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Trans. Signal Process.*, 46(2):375–390, Feb 1998.

[39] T. E. Bishop, D. Babacan, B. Amizic, T. Chan, R. Molina, and A. K. Katsaggelos. *Blind image deconvolution: problem formulation and existing approaches*, chapter 1, pages 1–42. CRC Press, 2007.

[40] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *Proc. IEEE Comp. Vis. and Pat. Recog.*, pages 1964–1971, 2009.

[41] E. S. Meinel. Origins of linear and nonlinear recursive restoration algorithms. *J. Opt. Soc. Am. A*, 3(6):787–799, 1986.

[42] A. S. Carasso. The APEX method in image sharpening and the use of low exponent lévy stable laws. *SIAM J. Appl. Math.*, 63(2):593–618, 2003.

[43] Y. Hiraoka, J. W. Sedat, and D. A. Agard. Determination of three-dimensional imaging properties of a light microscope system. *Biophys. J.*, 57:325–333, February 1990.

[44] Z. Kam, D. A. Agard, and J. W. Sedat. Three-dimensional microscopy in thick biological samples: A fresh approach for adjusting focus and correcting spherical aberration. *Bioimaging*, 5(1):40–49, 1997.

[45] L. Sherman, J. Y. Ye, O. Albert, and T. B. Norris. Adaptive correction of depth-induced aberrations in multiphoton scanning microscopy using a deformable mirror. *J. Microsc.*, 206(1):65–71, 2002.

[46] M. J. Booth, M. A. Neil, R. Juskaitis, and T. Wilson. Adaptive aberration correction in a confocal microscope. *Proc. Natl. Acad. Sci.*, 99(9):5788–5792, 2002.

[47] Z. Kam, P. Kner, D. Agard, and J. W. Sedat. Modelling the application of adaptive optics to wide-field microscope live imaging. *J. Microsc.*, 226:33–42, 2007.

[48] Z. Kam, B. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat. Computational adaptive optics for live three-dimensional biological imaging. *Proc. Natl. Acad. Sci.*, 98(7):3790–3795, 2001.

[49] B. M. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat. Phase retrieval of widefield microscopy point spread functions. In J.-A. Conchello, C. J. Cogswell, and T. Wilson, editors, *Proc. SPIE*, volume 4261, pages 60–68. SPIE, 2001.

[50] M. J. Booth. Adaptive optics in microscopy. *Philos. Transact. A Math. Phys. Eng. Sci.*, 365(1861):2829–2843, December 2007.

[51] J. C. Dainty and M. A. Fiddy. The essential role of prior knowledge in phase retrieval. *J. Mod. Opt.*, 31(3):325–330, 1984.

[52] J. Nocedal and S. J. Wright. Quasi-Newton methods. In *Numerical Optimization*, chapter 8, pages 193–201. Springer-Verlag, New York, 1999.

[53] T. Peng, A. Balijepalli, S. K. Gupta, and T. LeBrun. Algorithms for on-line monitoring of micro spheres in an optical tweezers-based assembly cell. *Journal of Computing and Information Science in Engineering*, 7(4):330–338, 2007.

[54] T. D. Visser, J. L. Oud, and G. J. Brakenhoff. Refractive index and axial distance measurements in 3-D microscopy. *Optik*, 90:17–19, 1992.

[55] K. Kim and G. Shevlyakov. Why Gaussianity? *IEEE Signal Process. Mag.*, 25(2):102–113, March 2008.

[56] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399